UNIVERSIDADE FEDERAL DE SANTA CATARINA

Curso de Pós Graduação em Engenharia Mecânica

CAPACIDADE DOS SISTEMAS DE MEDIÇÃO PARA TAREFAS DE INSPEÇÃO 100%

Tese submetida à Universidade Federal de Santa Catarina para obtenção do grau de Doutor em Engenharia Mecânica

Autor: *Gustavo Daniel Donatelli*

Orientador: *Prof. Dr.-Ing. Carlos Alberto Schneider*

Florianópolis
Santa Catarina - BRASIL
Maio 1999

CAPACIDADE DOS SISTEMAS DE MEDIÇÃO PARA TAREFAS DE INSPEÇÃO 100%

Gustavo Daniel Donatelli

Esta tese foi julgada adequada para obtenção do título de

"DOUTOR EM ENGENHARIA"

Especialidade ENGENHARIA MECÂNICA – Área de concentração METROLOGIA E INSTRUMENTAÇÃO, e aprovada em sua forma final pelo CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Prof. Carlos Alberto Schneider, Dr.Ing.
ORIENTADOR

Júlio César Passos

Prof. Júlio César Passos, Dr.
COORDENADOR DO CURSO DE PÓS-GRADUAÇÃO

BANCA EXAMINADORA

Albertazzi

Prof. Armando Albertazzi Gonçalves Jr., Dr.Eng.

Prof. Paulo Augusto Cauchick Miguel, Ph.D.

R. Freudenberg

Prof. Reinhard Freudenberg, Dr.Ing.

Prof. Robert Wayne Samohyl, Ph.D.

Prof. Carlos Alberto Schneider, Dr.Ing.

A Liliana e Marcos, que me
apoiaram em todo momento.

A meus pais, que me
mostraram o caminho.

A Mark Kassilov Golfan,
que esteve no início, mas
teve que partir sem ver o
final.

Agradecimentos

Ao Prof. Carlos Alberto Schneider, pela confiança dispensada e pelo apoio irrestrito às minhas iniciativas.

A Mark Kassilov Golfan e Edson Augusto Giesel, bolsistas de iniciação cientifica que tanto fizeram para o êxito deste trabalho.

Aos colegas do Labmetro e do CMCQ, pela compreensão e pelas longas e proveitosas discussões que contribuíram ao meu crescimento.

À UNIVERSIDADE FEDERAL DE SANTA CATARINA, particularmente ao Labmetro, por receber-me e viabilizar de diversas formas este trabalho.

À Fundação CERTI, particularmente ao Centro de Metrologia e Controle da Qualidade, pelo suporte nestes cinco anos de trabalho duro.

À UNIVERSIDADE NACIONAL DEL COMAHUE, pela possibilidade concedida e pelo apoio econômico.

À CAPES, pela bolsa de estudo concedida e pelo suporte dado a este trabalho através de bolsas de iniciação científica.

Aos amigos brasileiros, pela calidez, e aos argentinos, por não abandonar-me neste cinco anos de afastamento.

Resumo

Nesta tese é tratado o problema da avaliação da capabilidade de sistemas de medição para tarefas de inspeção 100%. Depois de discutir os critérios e procedimentos disponíveis, é proposto um novo índice da qualidade de inspeção, denominado *D*. O índice proposto avalia a proporção da perda de qualidade adimensional e unitária que pode ser atribuída aos erros de inspeção. Ele é definido utilizando a função perda quadrática, modificada para satisfazer os requisitos da inspeção 100%. São fornecidas as equações para calcular o valor de *D* em diferentes casos, tais como inspeção de tolerâncias uni- e bi-laterais e classificação dimensional. Utilizam-se funções perda nominal-é-melhor, assimétrica e menor-é-melhor dependendo do caso.

Considerando as dificuldades de avaliar a qualidade de inspeção na indústria, é proposto o uso da simulação computacional. O algoritmo proposto é baseado na hipótese que os erros de inspeção são uma conseqüência dos erros de medição. Ele é aplicado para avaliar o comportamento do índice *D* numa ampla diversidade de condições de inspeção e também é usado como núcleo de um protótipo de software denominado *Wininspect*. Este programa foi projetado como ferramenta para o melhoramento da qualidade de produto quando a inspeção 100% está envolvida no processo. O seu uso permite otimizar os parâmetros da inspeção considerando: a perda total da qualidade, a fração dessa perda que pode atribuir-se ao sistema de medição e o grau de contaminação do lote aceito com unidades não conformes. Dois estudos de caso são descritos, com o intuito de demostrar a aplicação do programa de simulação no controle da qualidade industrial.

São apresentadas recomendações para orientar a seleção e aplicação de sistemas de medição em tarefas de inspeção 100%. Em particular, o problema de conseguir *zero-defeito* por inspeção é discutido, mostrando que o valor do deslocamento dos limites de aceitação é um ponto chave na negociação entre cliente e fornecedor. Mostra-se também que esta política de fabricação produz elevados custos de falha interna quando a inspeção 100% é usada para preservar a qualidade do produto.

FEDERAL UNIVERSITY OF SANTA CATARINA

Graduate Course in Mechanical Engineering

CAPABILITY OF MEASUREMENT SYSTEMS FOR 100% INSPECTION TASKS

Dissertation submitted to the Federal University of Santa Catarina to obtain the degree of Doctor in Mechanical Engineering

Author: *Gustavo Daniel Donatelli*

Advisor: *Prof. Dr.-Ing. Carlos Alberto Schneider*

Florianópolis
Santa Catarina - BRAZIL
May 1999

Abstract

In this thesis the problem of evaluating the capability of measurement systems for 100% inspection tasks is addressed. After discussing the available criteria and procedures, a new measure of the inspection performance is proposed. The measure, called *D*, evaluates the fraction of the non-dimensional quality loss per unit that can be attributed to inspection errors. It is defined using the quadratic quality loss function concept, modified to fit 100% inspection requirements. Equations to compute the value of *D* are presented for several inspection cases: two-sided tolerances, one-sided tolerances and dimensional classification. Nominal-the-best, asymmetric and smaller-the-better quality loss functions are used, depending on the inspection case.

Considering the difficulties of evaluating the inspection performance in industrial situations, the use of computer simulation is proposed. The proposed algorithm is based on the assumption that inspection errors are caused by the lack of measurement accuracy. It is applied to evaluate the behaviour of the *D*-measure under a broad set of inspection conditions and constitutes the core of a software prototype called *Wininspect*. This software prototype has been designed as a tool for quality improvement when 100% inspection is included in the process flowchart. Its use would permit optimising the inspection conditions regarding the total quality loss, the fraction of quality loss due to defective inspection and the degree of contamination of the accepted batch. Two case studies have been included to show the application of the simulation software in industrial quality control situations.

Based on simulation results, recommendations are drawn to guide the selection and application of measurement systems in 100% inspection tasks. In particular, the problem of achieving zero-defect by 100% inspection is discussed, showing that the size of limit displacements is a key issue in customer-supplier negotiation. It is also shown that this manufacturing policy results in high internal failure costs when 100% inspection is used to preserve the quality of the product.

INDEX

INDEX OF FIGURES

INDEX OF TABLES

LIST OF ACRONYMS

Symbol	Definition
1D, 2D, 3D	1-Dimensional, 2-Dimensional, 3-Dimensional.
SPC	Statistical Process Control.
PCA	Process Capability Analysis.
C	Conforming (unit).
NC	Non-conforming (unit).
QC	Quality Control (QC-system, QC-application, QC-process).
ISO	International Organization for Standardization.
GUM	ISO Guide for the Expression of Uncertainty in Measurement.
GPS	Geometrical Product Specifications.
PDF	Probability Density Function.
LSL	Lower Specification Limit.
USL	Upper Specification Limit.
QLF	Quality Loss Function.
GPC	Gage Performance Curve.
LAL	Lower Acceptance Limit.
UAL	Upper Acceptance Limit.
OLS	Least Squares Method.

1 BACKGROUND AND RESEARCH FOCUS

Market forces generated by the international competition have caused an increasing demand on product quality and a permanent pressure for the reduction of manufacturing costs. Manufactured quality characteristics must satisfy tolerances that are becoming more and more strict. This imposes heavy requirements on the capability of manufacturing processes, but also on the performance of quality control systems.

In this chapter several techniques used today to assure the quality of manufactured units are reviewed. The operational advantages of using measurement systems in quality control are shown and the need for 100% inspection in particular cases are presented and justified. After discussing the available criteria to assess the capability of production measurement systems, the focus of this thesis is defmed.

1.1 From the quality of the product to the quality of measurement results

The functional quality and production costs of a mechanical engineering product depend largely on the geometrical quality of its parts, specified in drawings by means of dimensions and tolerances. Tolerances are geometrical constraints that define 1D, 2D or 3D regions within which a manufactured surface, or some parameter derived from it, must be in order that the units can be considered acceptable.

The objective of any manufacturing operation is to produce parts that fulfil tolerance specifications. Several techniques are applied today to achieve this ideal, such as statistical process control (SPC), process capability analysis (PCA) and inspection. Statistical process control allows identifying the action of special causes of variation and provides the information for their elimination. Capability analysis is aimed to define whether a manufacturing process is capable of satisfying a given tolerance. From evidence of insufficient capability, actions to improve the process could be decided and implemented. The proper application of these two techniques can prevent the production of non-conforming units, so leading to an economic and efficient production. Because of that, SPC and PCA are the preferred tools in those companies that have adopted the continuous improvement approach as a manufacturing strategy.

Inspection can be defined as *"...the process of measuring, examining, testing, gaging, or otherwise comparing the unit (element) with applicable requirements"* /1/. It provides the basis for actions on the product, typically separation of non-conforming items. The exclusive use of this technique to assure product quality is not recommended, because of its lack of economic effectiveness. However, only inspection can provide conclusive evidence on whether each unit fulfils the specification or not. Because of this, 100% inspection is used as a complement of SPC and PCA in case of characteristics that can produce critical failures or for which the quality target is "zero-defects" /2/. It is also applied when the manufacturing process is not capable (e.g. $Cp \leq 1.33$) or to separate units into classes according to their size (e.g. classification of bearing elements, car pistons and other high precision products) /3/. In other cases, 100% inspection is required to satisfy legal or political requirements /4/. Finally, when automatic instruments are used, 100% inspection can be the most cost-effective approach to assure product quality /2/.

In engineering products, SPC, PCA and inspection can be performed on the basis of *attributes* or *variables*. Attributes describe only the status of a characteristic regarding the fulfilment of specification (e.g. conforming (C) or non-conforming (NC)). The most traditional technique to generate attributes is gaging, that can be defined as the physical comparison of a workpiece characteristic with gages embodying the specification limits. In figure 1.1 the functional-block model of a generic quality control (QC) operation based on 100% gaging is depicted. The attributes can be used to separate non-conforming units and then, if required, post-processed to allow SPC and PCA. Nevertheless, the information contained in attributes is not adequate for SPC purposes: if there is a process shift, its direction can not be known before rejection /5/. Because of this, the control of manufacturing processes based on attributes can not avoid the production of non-conforming units. The problem of information quality affects PCA as well: it is impossible to achieve a thorough description of process distribution on the basis of attributes. This leads to uncertain capability estimates, with consequences on the actions taken to maintain or improve the capability of the process.

On the other hand, attributes are adequate for inspection. However, the use of gaging to generate this type of data has well known operational drawbacks:

- inspection with fixed gages depends largely on skill and judgement of the user /6/;

- the need to provide wear allowances and manufacturing tolerances for gages results in a high rejection-rate of conforming units /7/;
- complex gages are expensive and often require too much time to be ready for service.

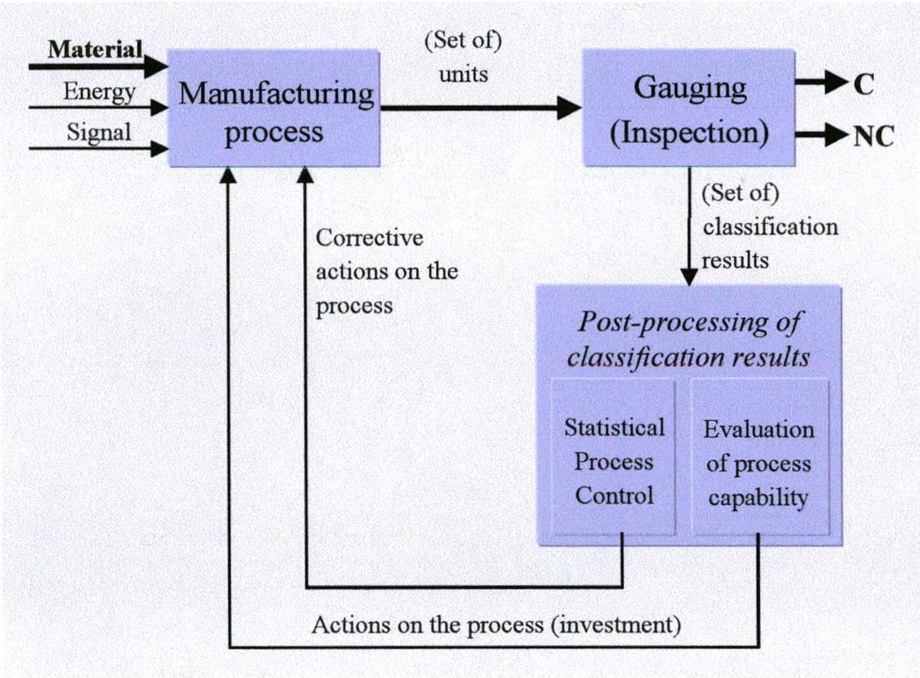

Figure 1.1: Gaging inspection in industrial quality control.

The problems above are minimised when the QC-process relies on variables, generated by measurement. Figure 1.2 shows a functional block model of a generic QC-operation based on measurement. Variables improve SPC and PCA functions, because the deviation from target is known quantitatively for each measured characteristic. Inspection quality is also improved: measuring instruments overcome the operational disadvantages of fixed gages. They are more flexible, allowing a cost-effective QC of medium and small production batches. The importance that industry gives to these features is shown by the distribution of investments between fixed gages and measuring instruments: today more than 90% of the investment in equipment for industrial metrology corresponds to measuring instruments /8/.

When variable data are used for inspection purposes, the conformity assessment is made comparing each measurement result with a set of specification limits, regardless whether the tolerance is a 1D-, 2D- or 3D-region. To assure the functional meaningfulness of this comparison, measurement results have to be consistent with the tolerance specification and also with the design intent /9/. It should be noted that conformity assessment converts variables into attributes, which are used to classify units in conforming or non-conforming.

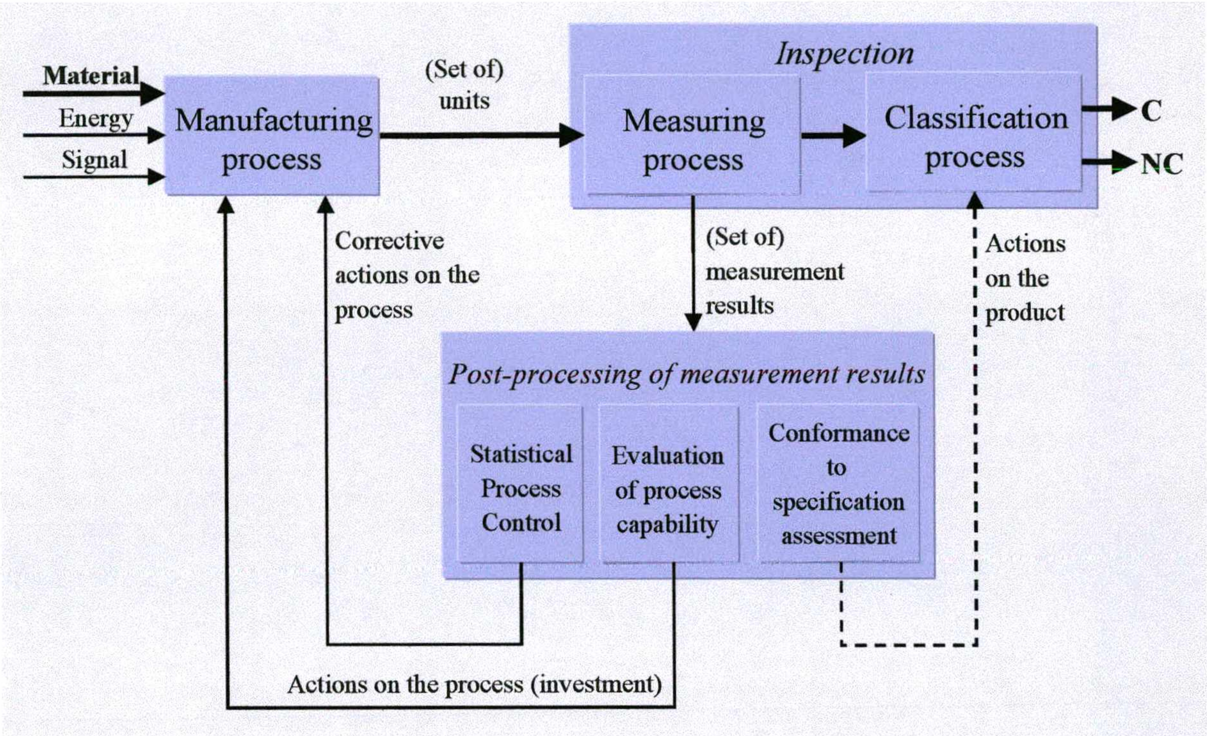

Figure 1.2: Measurement systems in industrial quality control.

However, just like fixed and functional gages, measuring instruments are far from ideal. The accuracy and precision of measurement results depends on several factors:

- characteristics of the instrument (contact point configuration, kinematic, elastic and dynamic behaviour, data reduction algorithms, rounding, wear, etc.);
- environment (temperature, humidity, vibrations, electric noise, etc.);
- operator (training, skill, motivation, etc.);
- workpiece (properties of material, deviations from perfect form, etc.).

The true effect of these influences is often not known, placing an uncertainty on the value of measurand that corresponds to a given measurement result (see /10/ for the definition of measurand). It is easy to realize that this uncertainty propagates to the post-processing stage (see figure 1.2) causing the actions on the process and on the product to depart from the ideal in an unknown value. Finally, these non-ideal actions deteriorate the quality of the product.

Due to this, to preserve the quality of manufactured products it is necessary evaluating first the quality of measurement results /11, 12/. The evaluation should be made in different stages of the system life cycle:

- during the design of the QC-facility;
- to accept the purchased/manufactured system and to set it free for service;
- to check its stability, which can be altered by normal wear or even by abuse.

The design of a QC-system is the group of activities leading to the complete definition of the instrument and all the operating conditions that determine its metrological behaviour /13/ (note that the selection of measuring devices is included in this definition). An effective design requires ability to predict the behaviour of the design object, to determine if it is adequate for further development. In preliminary design stages, this prediction has to be made without the aid of experimental data: only catalogue data and *a priori* knowledge are available. On the contrary, the second and third type of evaluation can be performed by experiment. The actual behaviour of the system can be measured and compared with specifications. The result of this comparison determines whether the QC-system behaves as intended. However, the production environment places other kinds of limitations on experimental-type assessment: economical ones. Particularly in the case of QC-systems that are already in service, the separation of the instrument from its task to perform the evaluation produces monetary losses due to production delays, product accumulation and/or duplication of equipment.

The presentation above suggests that the evaluation of QC-systems and the assessment of product quality are analogous processes: in both cases it is necessary to determine whether a quality characteristic is within some desirable limits. This analogy has three main consequences:

- There is a need to define indices of QC-system performance. These indices (or measures) should be meaningful from the viewpoint of the different applications of measurement systems in production.

- It seems adequate to apply the modern trends of quality assurance to achieve the desired quality of measurement results. Then, more efforts have to be made in preliminary stages of measurement system design, when performance data are weaker.
- The evaluation of QC-system performance is a technical and economical problem. It has to be solved as accurately as necessary, rather than as accurately as possible. In addition, time and cost of evaluation are relevant variables in the selection of the evaluation procedure.

These topics are briefly addressed in §1.2 and §1.3, to define and justify the scope of this thesis.

1.2 Criteria of assessment of measurement systems

It is affirmed that a measurement system is *capable* or has enough *capability* when a quality characteristic, selected to represent the metrological behaviour of the system, is within some acceptance interval. Then, the verification of a capability statement requires:

- the definition of a quality characteristic;
- the definition of an acceptance interval;
- a procedure to assign a value to the quality characteristic that corresponds to a given (real) measurement system.

The existing criteria for measurement system capability evaluation can be classified as shown in figure 1.3.

An accurate measurement is not the final objective of production metrology. The actual objective is to provide sound basis for actions on the process and/or actions on the product, which are intended to assure that only products of appropriate quality reach the hands of the customer. In this context, it seems natural to evaluate the performance of a measurement system by the accuracy of those actions (left branch in figure 1.3). For example, when the measurement results are used for SPC, PCA and inspection (or classification), the capability of a measurement system could be defined as:

$$\left\{\begin{matrix}\text{System is}\\ \text{capable}\end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix}\text{Capable}\\ \text{for SPC}\end{matrix}\right\} \wedge \left\{\begin{matrix}\text{Capable}\\ \text{for PCA}\end{matrix}\right\} \wedge \left\{\begin{matrix}\text{Capable for}\\ \text{inspection}\end{matrix}\right\} \tag{1}$$

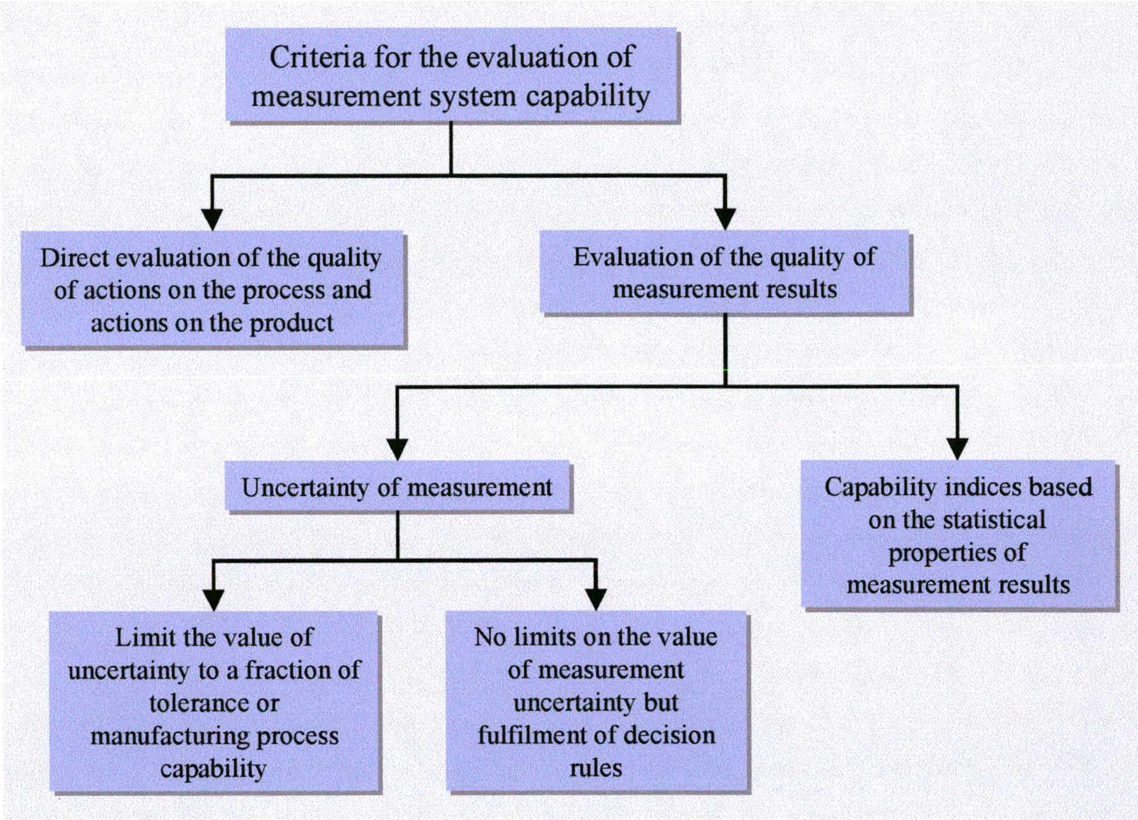

Figure 1.3: Existing criteria for the evaluation of measurement system capability.

Several investigations have been reported in this line of thought. Regarding the SPC application of measurement systems, Tricker *et al.* /14/ have shown that indicating device resolution or data rounding has an adverse effect on the performance of the range chart. It has been also shown that stochastic measurement errors reduce the sensitivity of the average and standard deviation chart ($\overline{X} - S$) with regard to process disturbances /15/.

The estimation of manufacturing process capability is also affected by measurement errors. From the statistical viewpoint, Mittag /16/ has studied the separate influence of random and constant measurement errors on the values of several capability indices. Random errors increase the spread of the distribution of measurement results with respect to the distribution of *true* values, leading to conservative estimates of process capability. Constant systematic errors produce a shift in the distribution mean, distorting the estimates of capability indices like Cpk, Cpm and Cpmk. Recognising this undesirable behaviour, several authors proposed equations to correct the values of indices estimated in the shop floor. Hernla /17/ and

Weckenmann /18/ have proposed to estimate the *true* variance of manufactured dimensions reducing the variance of measurement results by the square of standard measurement uncertainty. On the other hand, Donatelli, Barp and Schneider /19/ have suggested that systematic contributions to uncertainty should not be treated as independent random variables but as random curves. Under this assumption, a 95% confidence interval can be computed for the *true* values of process capability indices. For the sake of simplicity, these authors have proposed to use the worst capability that is consistent with measurement condition, that is, the lower value of the 95% confidence interval.

The accuracy of inspection is also affected by measurement errors. In the QC of batches manufactured to fulfil one- or two-sided specifications, a fraction of the inspected sample results misclassified. Conforming parts are reported by the inspection system as being non-conforming and *vice versa*. A survey of the relevancy of these inspection errors and a statistical analysis of their effect on sampling acceptance can be found in /20/. Contamination of the accepted batch with non-conforming items is particularly against the main objective of 100% inspection, which is to assure zero-defects. This contamination can be avoided by the displacement of acceptance limits with respect to specification limits. However, there is a cost to be paid for this decision: if it is impossible to improve the capability of the manufacturing process, a number of conforming parts will be rejected by mistake. As mentioned above, another case of inspection is dimensional classification. The number of inspection errors depends mainly on the position of class limits with respect to manufacturing process distribution /21/: the higher the probability density of manufactured dimensions, the higher the number of classification errors. Because of this, dimensional classification imposes a particularly strict requirement on the accuracy and precision of measurement systems.

Several authors have studied the problem of assessing the performance of inspection systems. Most investigations are focused on gaging systems used to inspect one- and two-sided specifications. In spite of that, their results are also valid for measurement systems. The proposed measures depend on the probability of incorrect classification, thus evaluating the inspection performance at the level of actions on the product. They are described and discussed in chapter 2, along with the corresponding evaluation procedures.

The summary above makes evident that the lack of measurement accuracy is accepted to be a cause of the inaccuracy of actions on the process and actions on the product. This causality relationship is implicit in the second criterion for measurement system assessment. It establishes that the capability of a measurement system for a given QC-task can be evaluated through the quality of measurement results (right branch in figure 1.3).

Measurement uncertainty must be viewed as a statement of ignorance on the value of the measurand /22/. In the Guide to the Expression of the Uncertainty in Measurement (GUM), it is defined as "*...a parameter, associated with the result of a measurement that characterises the dispersion of the values that could reasonably be attributed to the measurand*" /23/.

Concepts like *true value* and *error* have not been used in the definition above. These quantities are not known and not knowable in the real world: they are theoretical constructs. This makes the definition in the GUM not consistent with other definitions of measurement uncertainty found in previous bibliography (see /24/, for example). Nevertheless, if the concept of error is accepted in a somewhat loose manner, it can be shown that measurement uncertainty determines probable bounds for measurement error. This is recognised in NIS 3003 /25/, that alternatively defines measurement uncertainty as "*...the range about zero in which the (measurement) error is thought to lie*". However, uncertainty should not be confused with a confidence interval for random errors. Indeed, most of the contributions to measurement uncertainty are of a systematic nature. /9, 26/.

According to this criterion, the capability of a measurement system is defined by:

$$\{\text{System is capable}\} \Leftrightarrow \{\text{Estimated uncertainty} \leq \text{target uncertainty}\} \qquad (2)$$

Without doubt this statement is simpler to verify in industrial practice than that expressed in (1). However, it requires defining the value of the target uncertainty, which is not a trivial task. When is uncertainty small enough?

Two criteria are used to define target uncertainty. One possibility is to limit the fraction of tolerance that can be consumed by measurement uncertainty. This is the case of the uncertainty per tolerance ratio U_{95}/T, which has been known for many years as the *golden rule of metrology* /17, 27/. Suggested maximum value of the ratio is 1/10, but measurement systems presenting ratios of 1/4 are common in industrial metrology.

The second criterion is the base of ISO/FDIS 14253-1 /28/ and ISO/DTR 14253-2(E) /29/, pertaining to the ISO-GPS chain of standards. The first document describes the decision rules to be used in customer-supplier relationships for proving conformance or non-conformance with specifications. To prove conformance with specifications, the supplier has to displace the specification limits by the value of expanded measurement uncertainty. As this technique is supposed to avoid the acceptance of non-conforming units, no limits are placed on the value of measurement uncertainty. The only constraint comes from the manufacturing economy. To avoid high scrap-rates the natural capability of the manufacturing process has to be reduced to fit in the acceptance interval. Therefore, the higher the measurement uncertainty the smaller the room for manufacturing process variation and *vice versa*. The companion document ISO/DTR 14253-2(E) establishes a simplified, iterative procedure to evaluate the uncertainty in industrial measurement and low-level calibrations. In this document the concept of target uncertainty is clearly defined and used as in expression (2).

Besides the capability criteria based on measurement uncertainty, other criteria exist that evaluate the capability of a measurement system using indices and procedures *ad hoc*. These have been proposed by companies and associations of manufacturers, generally related to the car industry (e.g. see /30, 31, 32/). The indices combine in different ways the statistical properties of measurement results, typically the standard deviation and the deviation of the mean of measurement results with respect to the value of a standard or master part. The capability statement is described by the following expression:

$$\{\text{System is capable}\} \Leftrightarrow \{\text{Estimate of capability index} \leq \text{Maximum recommended value}\} \quad (3)$$

Most recommendations differ in the procedure by which the statistical properties are estimated and in the equations to compute capability indices from the values of such properties. Therefore, when applied to the same instrument and measuring condition, they could lead to different results of the capability evaluation /33/. Because of this, the trend today is to use measurement uncertainty as the only index of measurement quality and measurement system adequacy /32, 34/.

It should be noted that the fulfilment of a capability statement based on the quality of measurement results implies a limitation on the permissible value of measurement errors. If it is accepted that measurement is not an objective but the means to achieve the desired product quality, some questions arise:

- What kind of relationship exists between the allowed measurement errors and the resulting inaccuracies of actions on the process and actions on the product?
- How the later affect product quality?
- Does the relationship depend on the statistical properties of the manufactured parts or not?
- How can the deterioration in product quality due to measurement inaccuracy be measured?

These questions are not completely answered in the literature. Thus, the evaluation of measurement system capability is still in the domain of the educated guess.

1.3 This thesis

This thesis addresses the problem of evaluating the performance of measurement systems dedicated to 100% inspection tasks. It should be considered a part of a major plan: the design of a set of measures and procedures to assess the capability of measurement systems with reference to product quality.

The main objectives are:

- to propose a measure (index) of inspection performance based on the effect of measurement inaccuracy on the quality of the product;
- to propose a methodology for the evaluation of the measure in practical situations;
- to study the behaviour of the measure and correlate it with the behaviour of uncertainty of measurement and other indices under the same conditions;
- to propose a solution for the evaluation of 100% inspection systems in industry.

In Chapter 2 the current measures of inspection performance are discussed, to set the basis for the formulation of a new measure in Chapter 3. Chapter 4 deals with the proposed evaluation procedure, which is computer simulation. It describes the mathematical models used to compute the value of the proposed measure from data available in the shop floor. Chapter 5 presents the results of evaluation of the measure behaviour for a broad set of inspection conditions. The relationships with other measures are also discussed. Chapter 6 deals with the description of a simulation package for industrial assessment and optimisation of inspection systems. Finally, in Chapter 7, the findings are summarised and some proposals for future works listed.

2 INSPECTION SYSTEM EVALUATION: STATE OF ART

In this chapter several measures of inspection performance are described. All of them evaluate directly the performance of inspection systems at the level of actions on the product, so they correspond to the left branch in the block diagram of figure 1.3. However, they differ from one another by the procedure of evaluation.

In §2.1 twelve measures of inspection performance are described and discussed. These measures have been developed for gaging systems. In spite of that, they can be also applied to measurement-based systems. A procedure for direct experimental evaluation is described.

The measures in §2.2 are conceptually and mathematically related to those in §2.1. The basic elements are the same, but they permit a more detailed analysis of the inspection performance in each specification limit. In addition, they are associated with a computational algorithm that derives the values of probabilities of inspection errors from the statistical description of manufacturing process and measurement errors. Thus, they are based on the assumption of causality discussed in §1.2.

In §2.3 a tool to predict the classification performance known as the *Gage Performance Curve* (GPC) is briefly presented. It should not be considered as a measure of inspection performance. The GPC is also based on the assumption of causality: it is drawn starting from the value of repeatability and reproducibility obtained by experiment.

2.1 Direct measures of inspection performance

Several measures have been proposed to evaluate the inspection performance of QC-systems. The concepts embodied in these measures are listed in table 2.1, together with their mathematical definitions. For more details on each particular measure and the statistical properties of their estimators, refer to the bibliography cited in the table. The measures μ_1, μ_2, μ_6, μ_7, μ_8 and μ_{11} are based on a philosophy that can be attributed to Juran, who suggested that the inspection performance should be evaluated independently of the incoming quality /35/. These measures combine the effect of two basic quantities: θ, probability of classifying a non-conforming unit as conforming and ϕ, probability of classifying a conforming unit as non-conforming.

Interpretation / References	**Definition**	**Eq.**
Probability of correct classification of non- conforming units /35/	$\mu_1 = 1-\theta$	(4)
Probability of incorrect classification of conforming units /35/	$\mu_2 = \phi$	(5)
Probability of correct classification /36, 37/	$\mu_3 = p\cdot(1-\phi)+q\cdot(1-\theta)$	(6)
Probability of a unit being conforming, given that it has been classified as conforming /38/	$\mu_4 = \dfrac{p\cdot(1-\phi)}{p\cdot(1-\phi)+q\cdot\theta}$	(7)
Probability of a unit being non-conforming, given that it has been classified as non-conforming /38/	$\mu_5 = \dfrac{q\cdot(1-\theta)}{p\cdot\phi+q\cdot(1-\theta)}$	(8)
Average of the differences between the probabilities of correct and incorrect classification for conforming and non-conforming units /39, 40/	$\mu_6 = 1-\phi-\theta$	(9)
Tiemstra's efficiency rate /41/	$\mu_7 = (1-\phi)\cdot(1-\theta)$	(10)
Average probability of correct classification for conforming and non-conforming units /42/	$\mu_8 = 1-\dfrac{(\phi+\theta)}{2}$	(11)
Change in error probability relative to the initial error probability /43/	$\mu_9 = (1-\theta)-\dfrac{p}{q}\cdot\phi$	(12)
Probability of correct decision, in excess of that due to chance, as a fraction of its maximum possible value /44/	$\mu_{10} = \dfrac{2\cdot p\cdot q\cdot(1-\phi-\theta)}{q+(1-2\cdot q)\cdot[p\cdot\phi+q\cdot(1-\theta)]}$	(13)
Odds ratio /45/	$\mu_{11} = \dfrac{(1-\phi)\cdot(1-\theta)}{\phi\cdot\theta}$	(14)
Correlation coefficient /45/	$\mu_{12} = \dfrac{p\cdot q\cdot(1-\phi-\theta)}{\sqrt{p\cdot q\cdot[p\cdot(1-\phi)+q\cdot\theta]\cdot[p\cdot\phi+q\cdot(1-\theta)]}}$	(15)

Table 2.1: Definition of twelve measures of inspection performance.

The remaining measures also depend on the value of q, the incoming fraction non-conforming, and $p = 1 - q$, the incoming fraction conforming (see μ_3, μ_4, μ_5, μ_9, μ_{10} and μ_{12}). In consequence, they reflect the interaction between inspection system characteristics and the condition under which it is used /46/.

The values of measures in table 2.1 can be estimated by experimentation. A sample of size n has to be inspected with the system to be evaluated and then re-inspected by a check-inspector provided with a master instrument. From the comparison of the results of both inspection processes, values of the following statistics can be obtained:

a → number of conforming units correctly classified

b → number of conforming units incorrectly classified

c → number of non-conforming units incorrectly classified

d → number of non-conforming units correctly classified

where n is the size of the sample:

$$n = a + b + c + d \tag{16}$$

From these values, p, q, θ, ϕ can be estimated as:

$$\hat{p} = (a + b)/n \tag{17}$$

$$\hat{q} = (c + d)/n \tag{18}$$

$$\hat{\theta} = c/(c + d) \tag{19}$$

$$\hat{\phi} = b/(a + b) \tag{20}$$

These four estimates can be replaced in equations (4)-(15) to compute the estimated measures of inspection performance, $\hat{\mu}_1$ to $\hat{\mu}_{12}$.

From the viewpoint of production metrology, some objections could be placed on the significance of the measures in table 2.1. First, the events of inspection error are not weighted by their actual influence on the quality of the outgoing product: they are just counted. Second, some of the measures do not make any difference between acceptance of non-conforming units and rejection of conforming ones (see equations of μ_6, μ_7, μ_8 and μ_{11}: the values of ϕ and

θ can be permuted without changing the value of the measure). Third, the measures do not distinguish the classification errors produced in the lower specification limits (LSL) from those produced in the upper specification limit (USL).

These drawbacks can be traced to the assumption of a default quality loss function (QLF) of the step-type (see figure 2.1). According to this model, the loss in quality due to the acceptance of a non-conforming unit does not depend on the difference between the value of measurand and the corresponding specification limit: it remains constant and equal to A_0, the cost of replacing or repairing a defective unit. The same reasoning can be applied in case of rejection of conforming units. This explains why the measures of the inspection performance reviewed in table 2.1 only depend on the number of inspection errors and not on their effect on the function.

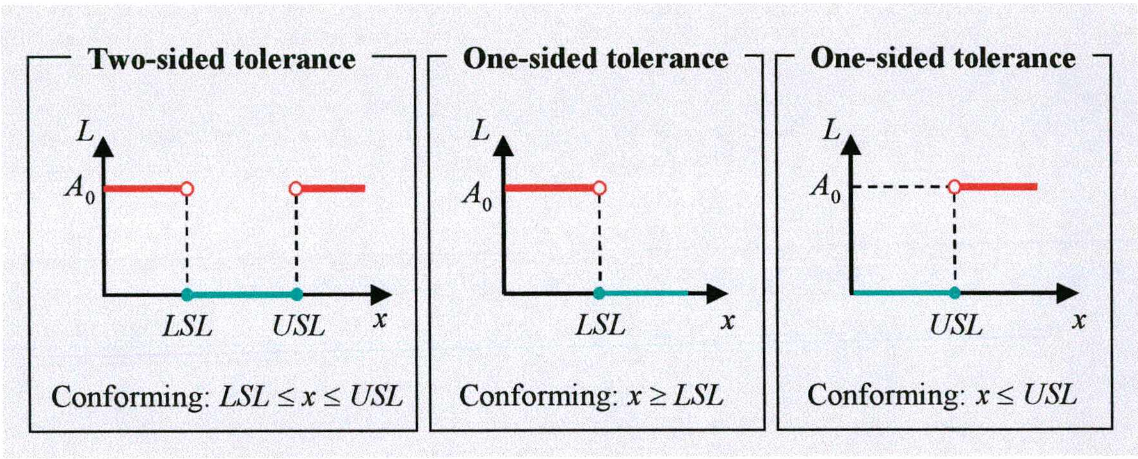

Figure 2.1: The step quality loss function for one- and two-sided tolerance specifications.

On the other hand, it has been suggested that measures μ_3 to μ_{12} could produce different results when ranking candidate inspection systems under certain particular conditions /46/. The results are consistent as long as one of the compared systems has both error probabilities (ϕ and θ) smaller than those of the other system. That is, for the comparison between two inspection systems A and B:

$$\phi_A < \phi_B \wedge \theta_A < \theta_B \Leftrightarrow \mu_{iA} > \mu_{iB} \qquad \forall i = 3 \ldots 12 \qquad (21)$$

If this condition is not fulfilled, the relative ranking provided by the different measures may vary: the result of the comparison becomes ambiguous.

In spite of these objections, it must be recognised that the measures and procedure above allow evaluating directly the inspection performance of QC-systems. The approach does not consider the cause of misclassification, but its effect on the inspection quality. This way, it could be applied to attribute gages as well as to measurement-based systems. However, two main issues should be considered to achieve a reliable experimental evaluation of the inspection performance:

- The size of the sample should be big enough to achieve a reliable estimation. No information has been found on this subject in the reviewed literature.
- The master inspection and the inspection made with the system to be evaluated have to be consistent, to prevent that within-part variations distort the number of errors.

These issues result in evaluation procedures that are both costly and time-consuming. On the other hand, if they were not respected, the assessment results would be rather not reliable.

2.2 Probabilities of error type I and type II

In previous works, Donatelli and Schneider /21, 47/ have addressed the problem of evaluating the inspection performance on a theoretical basis. In this case, inspection errors are considered as a consequence of measurement errors of random nature.

An inspection error event is produced when the lack of accuracy of the measurement system causes the *true* and *apparent* classification status of a unit to be different. The *true* classification status is determined by the relationship between the *true* value of the quality characteristic and the specification limit. The *apparent* classification status is determined by the relationship between the measured value and the acceptance limit. In table 2.2 the conditions for the existence of both types of inspection errors are detailed.

The practical meaning of these errors depends on the nature of the inspection operation and the position of the considered limit. When inspection is performed to separate parts in conforming and non-conforming, the acceptance of a non-conforming part in the lower specification limit (LSL) is an event of error type I. The same error type is produced by rejection of a conforming part in the USL. In the classification of parts for selective assembly, control errors produced in the inner class limits should be interpreted as events of misclassification.

Error Type	***True* classification status**	***Apparent* classification status**
I	*true* value < spec. limit	indication > acceptance limit
II	*true* value > spec. limit	indication < acceptance limit

Table 2.2: Definition of inspection errors type I and type II.

The probabilities of error type I and type II that can occur in the neighbourhood of a given specification limit are proposed to be indices to the quality of inspection:

$$P\{\text{Error type I}\}^{SLj} = \lim_{n\to\infty}\left(\frac{\text{number of errors type I}^{SLj}}{n}\right) \tag{22}$$

$$P\{\text{Error type II}\}^{SLj} = \lim_{n\to\infty}\left(\frac{\text{number of errors type II}^{SLj}}{n}\right) \tag{23}$$

where n is the total number of inspected units. In case of inspection of two-sided tolerances, the relationships among ϕ, θ and the probabilities of inspection errors are given by the following equations:

$$\phi = \frac{1}{p}\cdot\left[P\{\text{Error type II}\}^{LSL} + P\{\text{Error type I}\}^{USL}\right] \tag{24}$$

$$\theta = \frac{1}{q}\cdot\left[P\{\text{Error type I}\}^{LSL} + P\{\text{Error type II}\}^{USL}\right] \tag{25}$$

Note that the probabilities of error type I and II make possible evaluating the effect of measurement errors in dimensional classification operations. This would be impossible with ϕ, θ and the measures in §2.1.

A computational algorithm has been proposed to evaluate the measures of inspection performance in equations (22) and (23). The algorithm solves the numerical integration of the joint probability density function (PDF) of measurement and manufacturing errors within the domain of inspection errors in each specification limit (see details in /47/). In the proposed algorithm the beta PDF is adopted to represent the statistical behaviour of manufactured dimensions and measurement errors. Required data for the application in the assessment of real inspection facilities are:

- four parameters of a beta PDF characterising the variability of manufactured dimensions;
- four parameters of a beta PDF, characterising the random behaviour of measurement errors;
- position of the specification limits;
- position of the acceptance limits.

The scope of application of the proposed analytic procedure is constrained by the assumptions used in the construction of the mathematical model. These are:

- *true* dimensions of manufactured workpieces and measurement results are represented by continuous variables;
- all systematic errors have to be corrected prior to the use of measurement results in conformity assessment;
- the instrument is applied to each manufacturing unit once, being the classification result a consequence of that elementary measurement operation;
- the comparison between the measurement result and each one of the control limits is done on a numerical basis. It is considered a mathematical operation, with negligible processing errors.

The probabilities of inspection errors type I and type II evaluate the effect of measurement errors on the quality of actions on the product. The independent analysis in each specification limit overcomes one of the disadvantages of the measures in table 2.1: the reduction of information content. Unfortunately, the comparison of the inspection performance of several candidate systems becomes more complex than in the previous case.

It should be noted that probabilities of inspection errors are conceptually and mathematically related to the measures of inspection performance in table 2.1. On the one hand, both assume the default step quality loss function depicted in figure 2.1. On the other hand, the knowledge of the values of p and q along with equations (24) and (25) permit computing the measures in table 2.1 from the values of probabilities of inspection error defined in table 2.2.

The criticisms against the proposed methodology are related to the simplifications embodied in the mathematical model. In particular, the representation by continuous variables and the hypothesis of complete correction of systematic errors produce a heavy restriction to the application of the algorithm for the assessment of industrial inspection systems. In spite of these problems, the analytic computation of the probabilities of classification errors provides a valid approach for the *a priori* evaluation of inspection performance and also a low-cost approximate evaluation of existing systems.

2.3 Gage Performance Curve

The *Gage Performance Curve* has been proposed in the automotive industry to describe the classification behaviour of measurement systems using only repeatability and reproducibility data (*GR&R*) /31/.

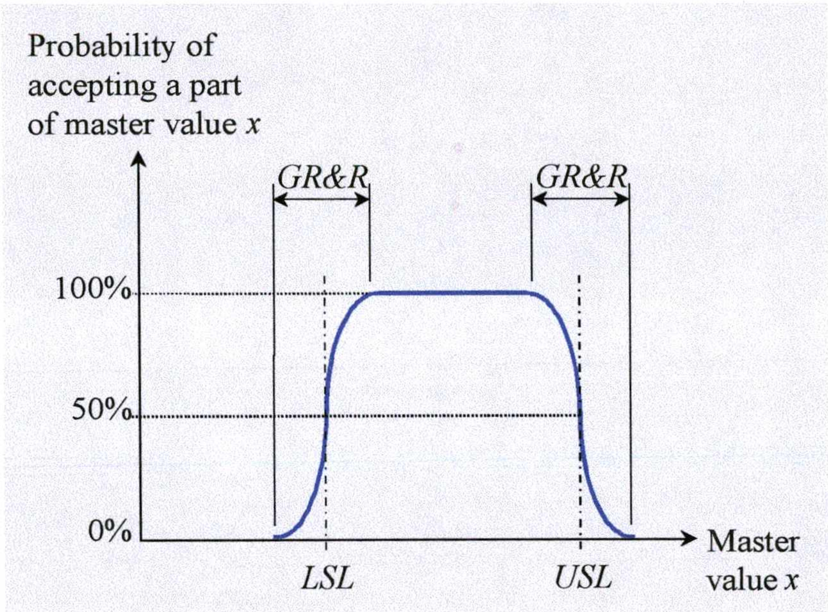

Figure 2.2: The Gage Performance Curve /31/.

As shown in the figure above, the GPC represents the probability of accepting a part of some master value and provides a basis for control limit displacement (figure 2.2).

The main objection to the use of the GPC is that it does not inform the number of inspection errors that can be expected for a given application. Indeed, the GPC does not depend on the statistical properties of the manufacturing process under inspection. So, it is neither related to those measures in §2.1 and §2.2, nor has any relationship with the losses in product quality due to misclassification.

However, the GPC has important consequences for the logic of this thesis: it reinforces the hypothesis of causality and suggests that, if measurement errors are mainly random, most of the misclassified units will have master values that are close to the specification limits.

2.4 On the properties of a good measure of inspection performance

In sections 2.1, 2.2 and 2.3 several kinds of measures of inspection performance have been described and discussed, along with their corresponding evaluation procedures. Some conclusions can be drawn from the presented information, regarding the properties that a measure of inspection performance should have to be useful in production metrology:

- The deterioration in product quality does not seem to be proportional to the number of inspection errors. If random measurement errors dominate, most inspection errors correspond to units with *true* values near the specification limits. A measure of inspection performance should weigh inspection errors by their effect on quality.
- Input quantities required for the evaluation of the inspection performance have to be available in the shop floor.
- The measure should not be tied to an experimental evaluation procedure; otherwise the *a priori* assessment becomes impossible.
- During inspection system design, it is necessary to determine whether the measurement system has enough capability for the inspection task. If the system were not capable, the design should be modified and tested again. This cycle has to be repeated until the inspection system design fulfils the requirements. Thus, the evaluation procedure should be as fast as possible.

These properties have been embodied in the measure and evaluation procedure proposed in the chapters 3 and 4 respectively.

3 A NEW MEASURE OF INSPECTION PERFORMANCE

Inspection is an industrial operation aimed to preserve the quality of delivered units. From this viewpoint, a completely accurate measurement will satisfy the objective. If measurement uncertainty is zero, the resolution of the indicating device is zero and all systematic errors have been corrected, the accepted batch will not be contaminated with non-conforming units. However, this is a somewhat idealistic situation: such a measurement system does not exist.

Industry is profit-driven, so that the selection of a measurement system should be made to minimise the overall quality costs. Internal and external failure costs increase as the inspection departs from ideal. On the other hand, the need to reduce inspection errors makes the appraisal costs to grow. A nearly optimum solution could be achieved if the necessary information is available. The contributors to the second cost group can be estimated without any problem. Equipment and installation costs, inspection times, manpower costs, calibration and validation costs are available in the company or can be obtained from the manufacturer of the instrument. The problem arises when the internal and external failure costs due to non-ideal inspection have to be estimated. It is evident that the value of measurement uncertainty will not provide the required information. Some of the direct measures of inspection performance in §2.1 and §2.2 are closer to what is needed. However, they are based on the step QLF, which is not a realistic quality loss model. Indeed, these measures operate at the level of actions on the product: they just count inspection errors. There is no clear relationship between the values of those measures and the quality loss due to non-ideal inspection.

In this chapter an economic measure of inspection performance is proposed. It operates directly at the level of internal and external failure costs of the manufactured batch, evaluating the effect of units accepted and rejected by mistake. Then, it provides the information needed to select or design a cost-effective inspection system.

3.1 Definition of the *D*-measure

The quadratic QLF, firstly introduced by Taguchi /48/, is adopted in this thesis to represent the behaviour of manufactured quality characteristics. According to this model, a characteristic whose value deviates from the functional target contributes to the total quality loss in a quantity that is proportional to the square of its deviation from target. Depending on the tolerance specification and the position of the functional target with respect to the specification limits, four types of quadratic loss functions can be distinguished: asymmetric, nominal-the-best, smaller-the-better and bigger-the-better. The variations of these functions for 100% inspected batches are depicted in figure 3.1.

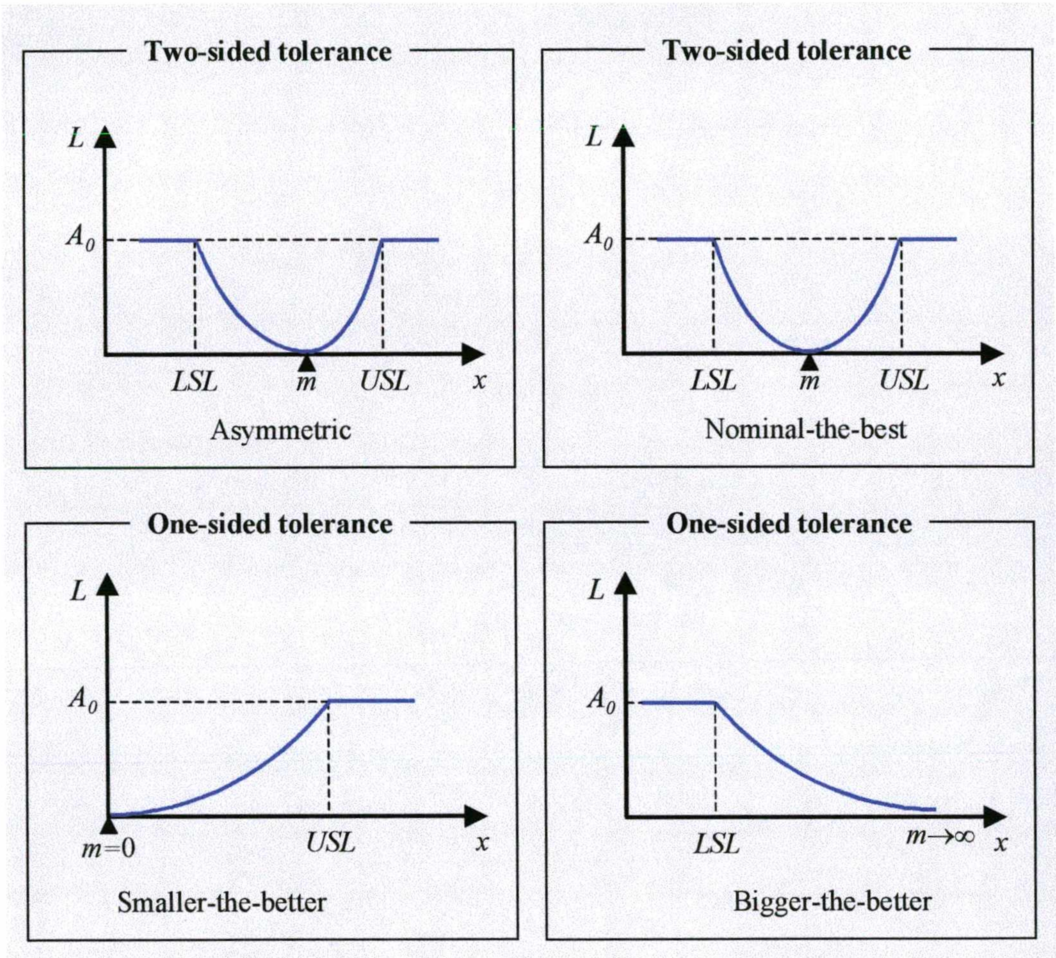

Figure 3.1: The quadratic loss function for one- and two-sided tolerance specifications.

It should be noted that, in case of two-sided tolerances, the cost of replacing or repairing a defective unit A_0 is assumed to be equal for units whose values are smaller than LSL and bigger than the USL /49/.

The functions in figure 3.1 show that units whose values are in the neighbourhood of a specification limit produce quality losses of the same order, regardless of their conformity status. This is the case of inspection errors: non-conforming parts, whose values are mostly status. This is the case of inspection errors: non-conforming parts, whose values are mostly close to specification limits, are classified by the inspection system as conforming and *vice versa*. It becomes evident that the quality loss due to an inspection error grows as the value of the misclassified unit departs from specification limit.

The proposal of the author is to measure the inspection performance in terms of the effect of classification errors on the final quality of the batch. The proposed index is referred to here as D-measure. Its mathematical definition is:

$$D = 1 - \frac{L^i}{L^r} \tag{26}$$

where L^i is the total quality loss that can be expected in a batch subjected to ideal inspection and L^r is the total quality loss in the same batch inspected with the system to be evaluated. Then, the D-measure can be defined as *the fraction of the total quality loss that can be attributed to inspection errors*.

It should be remembered that the quality loss is a property of the *true* value of a quality characteristic. This value can not be known in practical situations. However, a reasonable approximation can be obtained measuring the quality characteristics with a master instrument, having negligible uncertainty. This *master* value can be used to compute L^i and L^r. Thus, the quality loss L^r should not be confused with the *apparent* quality loss, that is a function of the deviation of the *measured* value with respect to the target.

The equations to compute the D-measure are derived in the following sections. The selected combinations of inspection task and quality loss function can be observed in table 3.1.

Inspection case	Quality loss function	Section
two-sided tolerances	asymmetric	§3.1.1
	nominal-the best	§3.1.2
one-sided tolerances	smaller-the-better	§3.1.3
dimensional classification	asymmetric	§3.1.4
	nominal-the best	§3.1.5

Table 3.1: Combinations of inspection tasks and quality loss functions.

3.1.1 Two-sided tolerances and asymmetric QLF

According to Tannock and Earl /50/, the total quality loss can be equated to internal and external failure costs. The use of an ideal measurement instrument in 100% inspection will assure that all defective units are segregated and do not reach the hands of the customer. In consequence, no external failure costs will be associated with those non-conforming units. However, each rejected unit contributes to the internal failure cost with a loss equal to A_0, the cost of replacing or repairing a defective unit. Then, the quality loss function $L^i(x)$ can be defined by the following set of equations (the supra-index i identifies ideal inspection):

$$L^i(x) = \begin{cases} k_l \cdot (x-m)^2 & \forall x \mid LSL \le x < m \\ k_u \cdot (x-m)^2 & \forall x \mid m \le x \le USL \\ A_0 & \text{elsewhere} \end{cases} \tag{27}$$

where x is the *true* value of the quality characteristic. The values of the lower-branch constant k_l and the upper-branch constant k_u can be obtained by equating the loss to A_0 in the specification limits:

$$\begin{aligned} k_l &= \frac{A_0}{(LSL-m)^2} \\ k_u &= \frac{A_0}{(USL-m)^2} \end{aligned} \tag{28}$$

Therefore, the total quality loss in a finite batch of size n can be expressed by:

$$L^i = A_0 \cdot \left[r^i + \sum_{j=1}^{a_l^i} \left(\frac{x_j - m}{LSL-m} \right)^2 + \sum_{k=1}^{a_u^i} \left(\frac{x_k - m}{USL-m} \right)^2 \right] \tag{29}$$

where r^i is the total number of non-conforming units, a_l^i is the number of conforming units with $x_j < m$ and a_u^i is the number of conforming units with $x_k \ge m$.

On the other hand, real measurement systems introduce errors. Therefore, a measurement result y departs from the corresponding value of the quality characteristic x:

$$y = x + \varepsilon \tag{30}$$

The decision on whether to accept a unit or not depends on the comparison of the measurement result with acceptance limits, *LAL* and *UAL*. These limits can be equal to specification limits or not, depending on the quality policy of the company. The classification is incorrect if any of following situations arise:

$x < LSL \wedge y \geq LAL \quad \Rightarrow \quad$ acceptance by mistake in LAL

$x \geq LSL \wedge y < LAL \quad \Rightarrow \quad$ rejection by mistake in LAL

$x \leq USL \wedge y > UAL \quad \Rightarrow \quad$ rejection by mistake in UAL

$x > USL \wedge y \leq UAL \quad \Rightarrow \quad$ acceptance by mistake in UAL

In these equations, quality characteristics whose values are equal to specification limits ($x = LSL$ and $x = USL$), are considered to fulfil the specification. In a similar way, measurement results equal to the acceptance limits ($y = LAL$ and $y = UAL$), determine the acceptance of the characteristic. These conditions are not mandatory, but they are common in industrial practice and will be used here to illustrate the nature of conformity assessment.

The effect of misclassification in the quality loss is shown in figure 3.2. Units accepted by mistake are sent to client though they do not fulfil the specification. In consequence, their quality loss increases from A_0 to the corresponding value on the parabola (see x_1 and x_4 in figure 3.2). Units that are rejected by mistake increase their quality loss from the corresponding value on the parabola to A_0 (see x_2 and x_3 in figure 3.2). In both cases, the loss increments δL_{x_j} grow as well as the value of the misclassified unit deviates from the specification limit.

If it is accepted that measurement errors are small with respect to the tolerance, it is clear that quality characteristics with $x < m$ can not generate measurement results $y > UAL$. The same can be affirmed for quality characteristics with $x \geq m$: they can not result in readings $y < LAL$. Then, a definition of the asymmetric quality loss function that includes effects of non-ideal inspection is:

$$L^r(x)=\begin{cases} k_l\cdot(x-m)^2 & \forall x \mid x<m\wedge y\geq LAL \\ k_u\cdot(x-m)^2 & \forall x \mid x\geq m\wedge y\leq UAL \\ A_0 & \forall x \mid y<LAL\vee y>UAL \end{cases} \tag{31}$$

Supra-index r indicates real inspection. It should be observed that the loss function still depends on the *true* value of quality characteristic. The measurement results are used only to decide if the loss has to be computed by the equation of the parabola or if it is equal to A_0.

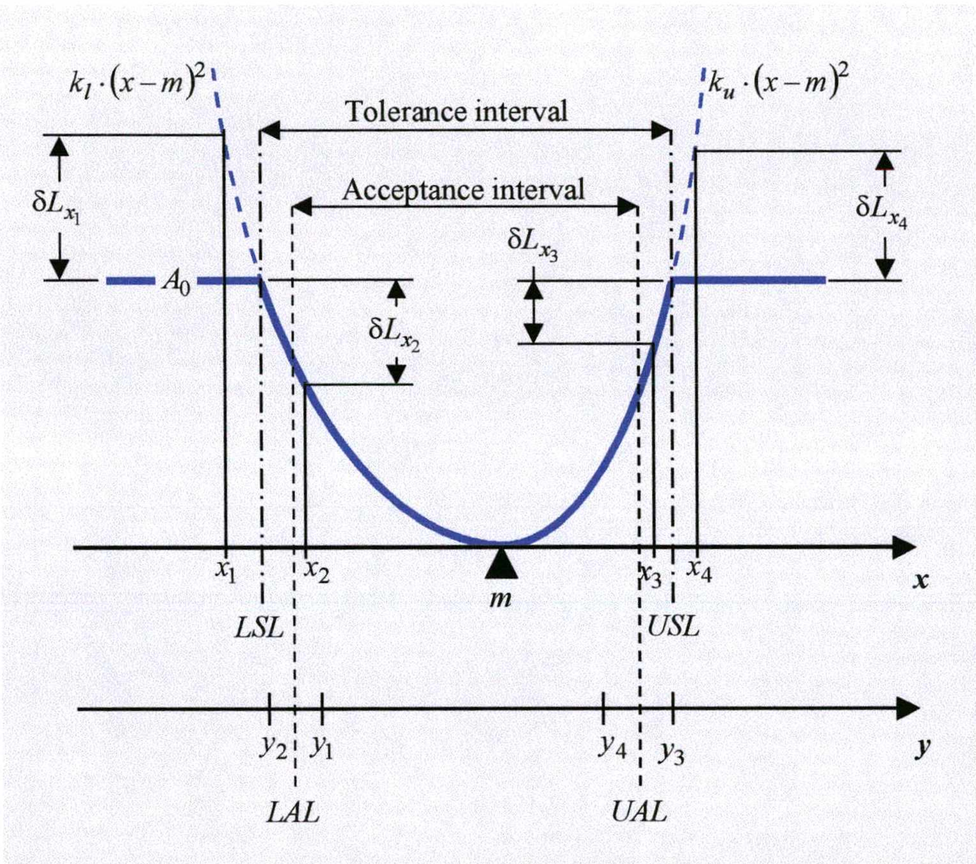

Figure 3.2: Incremental quality losses due to inspection errors (two-sided tolerances).

Thus, the total quality loss in a finite batch subjected to 100% inspection with a non-ideal measurement system can be expressed by:

$$L^r=A_0\cdot\left[r^r+\sum_{j=1}^{a_l^r}\left(\frac{x_j-m}{LSL-m}\right)^2+\sum_{k=1}^{a_u^r}\left(\frac{x_k-m}{USL-m}\right)^2\right] \tag{32}$$

where r^r is the number of units rejected by the inspection system and a_l^r and a_u^r are the number of accepted units, with *true* values $x_j < m$ and $x_k \geq m$ respectively. The total quality loss L^r can be attributed to a particular combination of manufacturing process, manufacturing target and specification limits, but also to classification errors.

For any unit of value x_j, the quality loss increment due to non-ideal inspection is:

$$\delta L_{x_j} = L^r(x_j) - L^i(x_j) \tag{33}$$

This incremental loss will be positive if the unit *j* has been misclassified and zero if it has been correctly classified by the inspection system. Extending the concept to the total number of units in the inspected batch:

$$\Delta L = L^r - L^i \tag{34}$$

The value of the *D*-measure can be computed by means of the following equation:

$$D = \frac{\Delta L}{L^r} = 1 - \frac{L^i}{L^r} \tag{35}$$

Thus, the *D*-measure defines a scale of inspection performance, in which *D*=0 corresponds to ideal inspection. An increase in the value of *D* has to be interpreted as a deterioration of inspection quality.

An alternative equation, which expresses the value of *D* in terms of non-dimensional quality losses per unit $\lambda^i = L^i/(A_0 \cdot n)$ and $\lambda^r = L^r/(A_0 \cdot n)$ is as follows:

$$D = 1 - \frac{\lambda^i}{\lambda^r} \tag{36}$$

where λ^i and λ^r can be computed by:

$$\lambda^i = \frac{r^i}{n} + \frac{1}{n} \cdot \left[\sum_{j=1}^{a_l^i} \left(\frac{x_j - m}{LSL - m} \right)^2 + \sum_{k=1}^{a_u^i} \left(\frac{x_k - m}{USL - m} \right)^2 \right] \tag{37}$$

$$\lambda^r = \frac{r^r}{n} + \frac{1}{n} \cdot \left[\sum_{j=1}^{a_l^r} \left(\frac{x_j - m}{LSL - m} \right)^2 + \sum_{k=1}^{a_u^r} \left(\frac{x_k - m}{USL - m} \right)^2 \right] \tag{38}$$

Applying equations (36), (37) and (38), the value of the D-measure can be estimated without the knowledge of the cost of replacing or repairing a product unit.

3.1.2 Two-sided tolerances and nominal-the-best QLF

The equations for the nominal-the-best QLF can be obtained from those derived in §3.1.1, considering that the functional target is in the middle of the tolerance interval:

$$m = \frac{USL + LSL}{2} \tag{39}$$

The expressions of the non-dimensional quality losses per unit are as follows:

$$\lambda^i = \frac{r^i}{n} + \frac{4}{n} \cdot \sum_{j=1}^{a^i} \left(\frac{x_j - m}{USL - LSL} \right)^2 \tag{40}$$

$$\lambda^r = \frac{r^r}{n} + \frac{4}{n} \cdot \sum_{j=1}^{a^r} \left(\frac{x_j - m}{USL - LSL} \right)^2 \tag{41}$$

where a^i is the total number of conforming units and a^r is the total number of units accepted by the inspection system. The other concepts remain unchanged, as well as the equation to compute the value of the D-measure from the non-dimensional quality losses above.

3.1.3 One-sided tolerances and smaller-the-better QLF

This particular combination of tolerance specification and quality loss function can be analysed using the same concepts developed for two-sided tolerances in §3.1.1. In case of ideal inspection, the smaller-the-better QLF is defined by:

$$L^i(x) = \begin{cases} \text{not defined} & x < 0 \\ k \cdot x^2 & \forall x \mid 0 \le x \le USL \\ A_0 & x > USL \end{cases} \tag{42}$$

where the constant k can be computed by:

$$k = \frac{A_0}{USL^2} \tag{43}$$

The total quality loss in a finite batch of size n is obtained by the sum of the cost of replacing or reworking defective units plus the internal and external failure costs produced by those units that are in tolerance but whose values depart from target:

$$L^i = A_0 \cdot \left[r^i + \sum_{j=1}^{a^i} \left(\frac{x_j}{USL} \right)^2 \right] \tag{44}$$

where r^i is the number of non-conforming units and a^i is the number of conforming ones, being:

$$n = r^i + a^i \tag{45}$$

The influence of non-ideal inspection on the quality loss is depicted in figure 3.3. Two kinds of inspection errors are possible:

- $x \leq USL \wedge y > UAL$ $\Rightarrow$ rejection by mistake in UAL
- $x > USL \wedge y \leq UAL$ $\Rightarrow$ acceptance by mistake in UAL

Then, the total quality loss in a finite batch inspected 100% with a non-ideal inspection system can be expressed as:

$$L^r = A_0 \cdot \left[r^r + \sum_{j=1}^{a^r} \left(\frac{x_j}{USL} \right)^2 \right] \tag{46}$$

where the supra-index r identifies real inspection and r^r and a^r are, respectively, the number of units rejected and accepted by the inspection system.

The non-dimensional quality losses per unit for ideal and real inspection are, respectively:

$$\lambda^i = \frac{r^i}{n} + \frac{1}{n} \cdot \sum_{j=1}^{a^i} \left(\frac{x_j}{USL} \right)^2 \tag{47}$$

$$\lambda^r = \frac{r^r}{n} + \frac{1}{n} \cdot \sum_{j=1}^{a^r} \left(\frac{x_j}{USL} \right)^2 \tag{48}$$

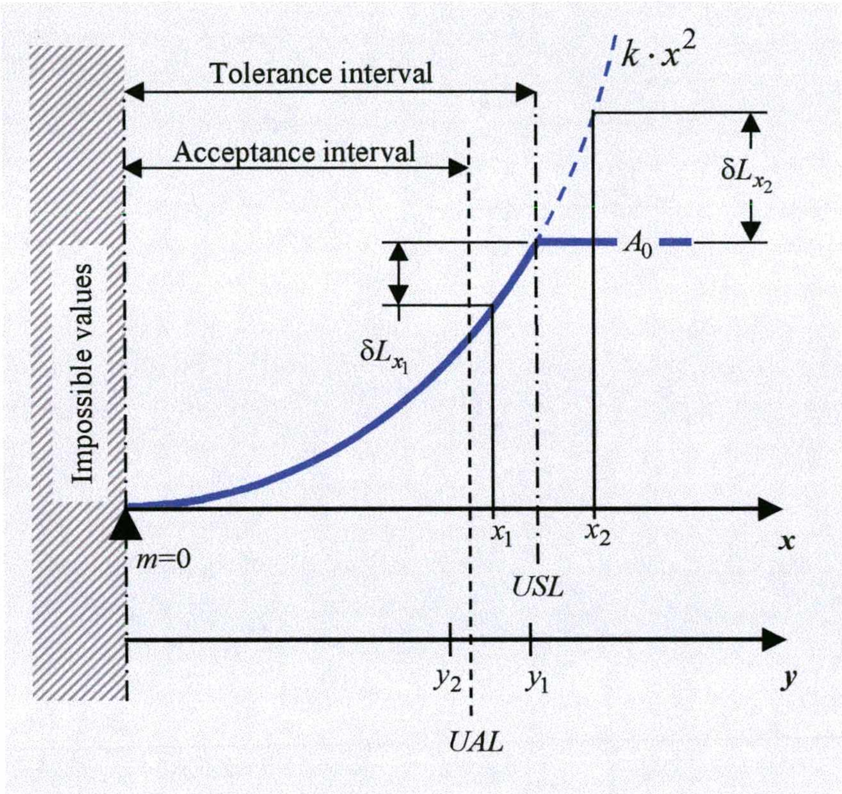

Figure 3.3: Incremental quality losses due to inspection errors (one-sided tolerances).

The value of the D-measure can be computed by means of equations (36), (47) and (48), as in previous cases.

3.1.4 Dimensional classification and asymmetric QLF

Classification, or sorting, is generally associated with the selective assembly manufacturing principle (for details on this principle and its properties, see Bjørke /51/). It is used in rolling bearing manufacturing and other high precision industries, which need to produce fits that are more precise than the processes available to manufacture individual parts. In this context, sorting is the separation of parts in a batch into several dimensional classes, up to ten in some cases. This is made by the comparison of the measurement results with the set of classification limits AL_j, that can be positioned in coincidence with the specification limits SL_j or not (see figure 3.4).

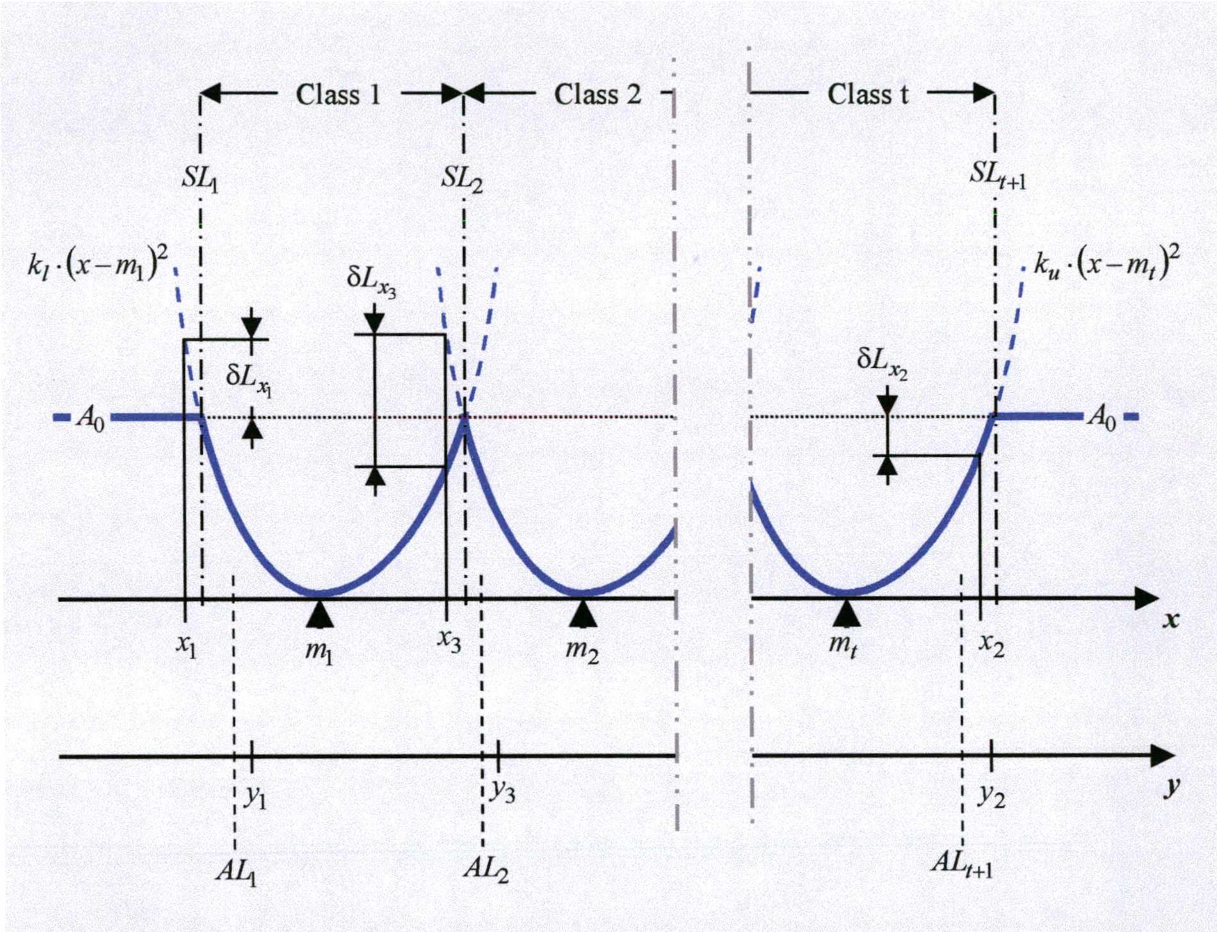

Figure 3.4: Incremental quality losses due to inspection errors (dimensional classification).

The quadratic QLF can be used within each dimensional class to compute quality losses due to dimensional variation. Quality characteristics whose *true* values are equal to the functional target of each class will produce zero loss in quality. On the other hand, quality characteristics whose values depart from the class target contribute to the quality loss with a failure cost that is proportional to the square of the deviation from target.

The assumptions for the study of the quality loss function are:

- all the classes have the same specification interval: $SL_{k+1} - SL_k = \text{constant} \ ; \forall k = 1 \ldots t$;

- the shape of the quality loss function remains constant for all classes: $k_{l_k} = k_l \wedge k_{u_k} = k_u$; $\forall k = 1..t$;
- in addition to t classes of conforming units, there exist two classes of non-conforming units defined by $x < SL_1$ and $x \geq SL_{t+1}$. Units in these classes are sent to scrap or rework, resulting in a quality loss per unit A_0 .

Under these assumptions, the QLF can be defined for the complete population classified with an ideal measurement system:

$$L^i(x) = \begin{cases} k_l \cdot (x - m_k)^2 & \forall x \mid SL_k \leq x < m_k \ ; \ k = 1...t \\ k_u \cdot (x - m_k)^2 & \forall x \mid m_k \leq x < SL_{k+1} \ ; \ k = 1...t \\ A_0 & \text{elsewhere} \end{cases} \tag{49}$$

The equation above can be used either to compute the total quality loss for each conforming class or a total quality loss in the complete classified batch. The first option is more consistent with the quality loss function philosophy. Nevertheless, this thesis deals with the variation of quality loss that can be attributed to non-ideal classification. So it is more adequate to use the second option:

$$L^i = A_0 \cdot \left\{ r^i + \sum_{k=1}^{t} \left[\sum_{p=1}^{a^i_{lk}} \left(\frac{x_p - m_k}{SL_k - m_k} \right)^2 + \sum_{q=1}^{a^i_{uk}} \left(\frac{x_q - m_k}{SL_{k+1} - m_k} \right)^2 \right] \right\} \tag{50}$$

where r^i is the total number of non-conforming units, a^i_{lk} and a^i_{uk} are respectively the number of conforming units with $x_p < m_k$ and $x_q \geq m_k$, belonging to the class k (the supra-index i identifies ideal classification).

Measurement errors result in classification errors that change the distribution of the units among the classes (see units x_1, x_2, x_3 in figure 3.4). If it is accepted that measurement errors are small if compared with the class interval, it is clear that quality characteristics with $x < m_k$ can not correspond to measurement results $y \geq AL_{k+1}$. The same can be affirmed for quality characteristics with $x \geq m_k$: they can not result in readings $y < AL_k$. Then, a definition of quadratic QLF that considers effects of real inspection is:

$$L^r(x) = \begin{cases} k_l \cdot (x - m_k)^2 & \forall x \mid x < m_k \wedge y \geq AL_k \ ; \ k = 1 \ldots t \\ k_u \cdot (x - m_k)^2 & \forall x \mid x \geq m_k \wedge y < AL_{k+1} \ ; \ k = 1 \ldots t \\ A_0 & \forall x \mid y < AL_1 \vee y \geq AL_{t+1} \end{cases} \tag{51}$$

Note that units whose measured values are equal to the class limits are accepted to belong to the class to the right of the limit. This condition is not mandatory, but has been arbitrary chosen here to develop the equations of the measure. The total quality loss in the batch classified by a non-ideal measurement system can be expressed as:

$$L^r = A_0 \cdot \left\{ r^r + \sum_{k=1}^{t} \left[\sum_{p=1}^{a_{l\,k}^r} \left(\frac{x_p - m_k}{SL_k - m_k} \right)^2 + \sum_{q=1}^{a_{u\,k}^r} \left(\frac{x_q - m_k}{SL_{k+1} - m_k} \right)^2 \right] \right\} \tag{52}$$

In this equation r^r is the number of units rejected by the inspection system and $a_{l\,k}^r$ and $a_{u\,k}^r$ are the number of units accepted in class *k*, with *true* values $x_p < m_k$ and $x_q \geq m_k$ respectively. Equations (50) and (52) can be transformed to apply the concept of non-dimensional quality loss per unit ($\lambda^i = L_t^i / A_0 \cdot n$ and $\lambda^r = L_t^r / A_0 \cdot n$):

$$\lambda^i = \frac{r^i}{n} + \frac{1}{n} \cdot \sum_{k=1}^{t} \left[\sum_{p=1}^{a_{l\,k}^i} \left(\frac{x_p - m_k}{SL_k - m_k} \right)^2 + \sum_{q=1}^{a_{u\,k}^i} \left(\frac{x_q - m_k}{SL_{k+1} - m_k} \right)^2 \right] \tag{53}$$

$$\lambda^r = \frac{r^r}{n} + \frac{1}{n} \cdot \sum_{k=1}^{t} \left[\sum_{p=1}^{a_{l\,k}^r} \left(\frac{x_p - m_k}{SL_k - m_k} \right)^2 + \sum_{q=1}^{a_{u\,k}^r} \left(\frac{x_q - m_k}{SL_{k+1} - m_k} \right)^2 \right] \tag{54}$$

Finally, the *D*-measure of the inspection performance can be computed using equations (53), (54) and (36).

3.1.5 Dimensional classification and nominal-the-best QLF

The use of the nominal-the-best QLF for the classes of conforming units does not result in relevant differences with respect to the case in §3.1.4. The target of each class is positioned at the same distance of the class limits, resulting in symmetric parabolas:

$$m_k = \frac{SL_k + SL_{k+1}}{2} \;;\; \forall k = 1..t \tag{55}$$

This allows simplifying the equations of the non-dimensional quality losses per unit:

$$\lambda^i = \frac{r^i}{n} + \frac{4}{n} \cdot \sum_{k=1}^{t} \left[\sum_{p=1}^{a_k^i} \left(\frac{x_p - m_k}{SL_{k+1} - SL_k} \right)^2 \right] \tag{56}$$

$$\lambda^r = \frac{r^r}{n} + \frac{4}{n} \cdot \sum_{k=1}^{t} \left[\sum_{p=1}^{a_k^r} \left(\frac{x_p - m_k}{SL_{k+1} - SL_k} \right)^2 \right] \tag{57}$$

where a_k^i is the total number of units belonging to the class k and a_k^r is the total number of units attributed to class k by the non-ideal measurement system. The other concepts remain unchanged, as well as the equation to compute the value of the *D*-measure from the non-dimensional quality losses above.

3.2 Combining the *D*-measure with the probabilities of inspection errors

In §3.1, equations of a new measure of inspection performance have been derived for some common cases in dimensional quality control. The proposed measure defines a consistent scale that permits comparing candidate measurement systems or assessing the adequacy of a measurement system for a given 100% inspection operation. The measure is sensitive to inspection errors, but weights their effect on internal and external failure costs by means of the quadratic QLF. It should be noted that no reference has been made to the causes of misclassification, but only to its effect on the final quality of the product, as desired. However, it can not be used for gaging systems, because its formulation requires knowledge of the values of quality characteristics used to estimate the measure.

In spite of its advantages, the D-measure has the same drawback of all summary measures: it hides information that could be useful to optimise inspection. When tuning a QC-system, it could be useful to know which type of inspection error makes the most important contribution to the incremental quality loss. It could also be an objective to avoid the acceptance of non-conforming units, regardless of the incremental quality loss introduced by the rejection of conforming units. Thus, it is suggested to use the D-measure together with probabilities of error type I and II, presented in §2.2. It should be remembered that these quantities are not consistent with the D-measure because they are associated with the step QLF. In consequence they should be considered an aid when the number of inspection errors is more important than their influence on the quality loss. This application will be exemplified in §6.2.

3.3 Estimating the value of the *D*-measure in practical situations

The procedures to estimate parameters of measurement systems can be classified into three groups: analytical, experimental and simulation. The value of the D-measure can be estimated by experiment, just as was described for other measures of inspection performance in table 2.1. A typical sequence of actions to perform this type of experiment could be:

a) Select a representative sample of product units;
b) a check inspector measures the sample units with a master instrument, compares the measurement results with the set of specification limits and classifies the units;
c) use the results of (b) to estimate the non-dimensional quality loss per unit $\hat{\lambda}^i$ (ideal inspection);
d) the operator measures the sample units with the instrument to be evaluated under real inspection conditions, compares the measured values with the set of acceptance limits and classifies the units;
e) use the results of (b) and (d) to estimate the non-dimensional quality loss per unit $\hat{\lambda}^r$ (real inspection);
f) Use the results of (c) and (e) to estimate the value of the D-measure: $\hat{D} = 1 - \hat{\lambda}^i / \hat{\lambda}^r$.

Some special care must be taken to achieve a realistic estimation of D-measure by experiment. First, the realisation of the measurand by the master measurement instrument has to be consistent with the realisation by the instrument to be evaluated. Second, if there is a lack of definition of the measurand leading to any intrinsic uncertainty contribution, it is necessary to measure the sample units in marked positions with both instruments. Third, if operator errors have to be evidenced, it is worth designing the experiment in such a way that the operator does not perceive that he or she is under evaluation. These conditions are difficult to satisfy in industrial environment (note that the second and third conditions are in conflict).

Even if it were possible to overcome the above-mentioned drawbacks, the experimental evaluation of the D-measure has several operational problems. It is time-consuming, because large samples are needed to achieve a reliable estimation of D. It is also error-prone, because it involves the computation of quality losses for each unit in the sample. Finally, it can not be applied during design of the inspection facility when only *a priori* knowledge is available.

If it is accepted that inspection errors are a consequence of measurement errors, the estimation of the D-measure can be made by analytic procedures or by simulation. Analytic procedures have to be discarded, because they limit the complexity of measurement error models to an unacceptable extent (see §2.2). On the contrary, simulation allows the use of more complicated models, providing a better representation of actual processes and does not require reference standards or instruments that form the basis of any metrological experiment /52/. In addition, it is time-efficient and does not require specially trained human beings once the computational algorithm is tuned.

In this thesis, it is proposed to address the evaluation of inspection performance by simulation. The basic idea is to emulate, by means of a computer program, the experiment described above. As the evaluation time is not a problem any more, the sample set can be as large as necessary to achieve a reliable estimation of the D-measure. Simulation results can be used to compute the probabilities of classification errors associated with each specification limit and other measures of inspection performance. Additionally, it becomes possible to predict the inspection performance from *a priori* knowledge of measurement errors, which is available during design of the QC-system.

The next chapter is dedicated to describe the simulated inspection algorithm. It will be used to analyse the behaviour of the D-measure and also to build the software for industrial evaluation of QC-systems.

4 A MODEL FOR SIMULATED INSPECTION

A simulation algorithm uses a mathematical model to emulate the behaviour of the system under analysis. The construction of this type of models involves the following actions /53/:

- identifying all the relevant variables;
- generating empirical assumptions about the relationships among the variables;
- introducing simplifications to allow the mathematical manipulation.

In the case of inspection performance evaluation, relevant input quantities come from three different sectors of the production activity: design, manufacturing and quality control (see figure 4.1). From this viewpoint, the evaluation procedure described in this chapter differs from other proposals of simulation in metrology, that rely on measurement system data alone (e.g. Megakal /54/). Design inputs are the specification limits and the functional target. The later is necessary to define the type of quality loss function used in the computation of *D*-measure. Manufacturing will inform the statistical properties of the process used to produce the part. The metrological behaviour of inspection system is defined by the contributions to measurement uncertainty. These bring into the simulation algorithm not only the characteristics of the instrument, but also the effects of environment. Note that acceptance limits have been included in the QC-process block and not in the design one. This is because limit displacements are a consequence of measurement uncertainty /28/. Indeed, if it were possible to perform an error-free measurement, the inspection could be made with respect to the specification limits with no risk of contaminating the accepted batch with non-conforming units.

The main empirical assumptions used to build the mathematical model of inspection are:

- inspection errors are caused by measurement errors;
- the comparison of measurement results with acceptance limits is error-free;
- measurement errors are small if compared with the tolerance.

On the other hand, the evaluation of the model by simulation has removed almost all the simplifying assumptions associated with analytic procedures. However, a fundamental restriction has to be introduced: coarse operator errors and other operator trends, like those producing flinching, are out of the scope of the model. The reason for this simplification is the difficulty to build a mathematical description of those behavioural patterns. Anyway, it does not affect the adequacy of the model for 100% automated inspection systems.

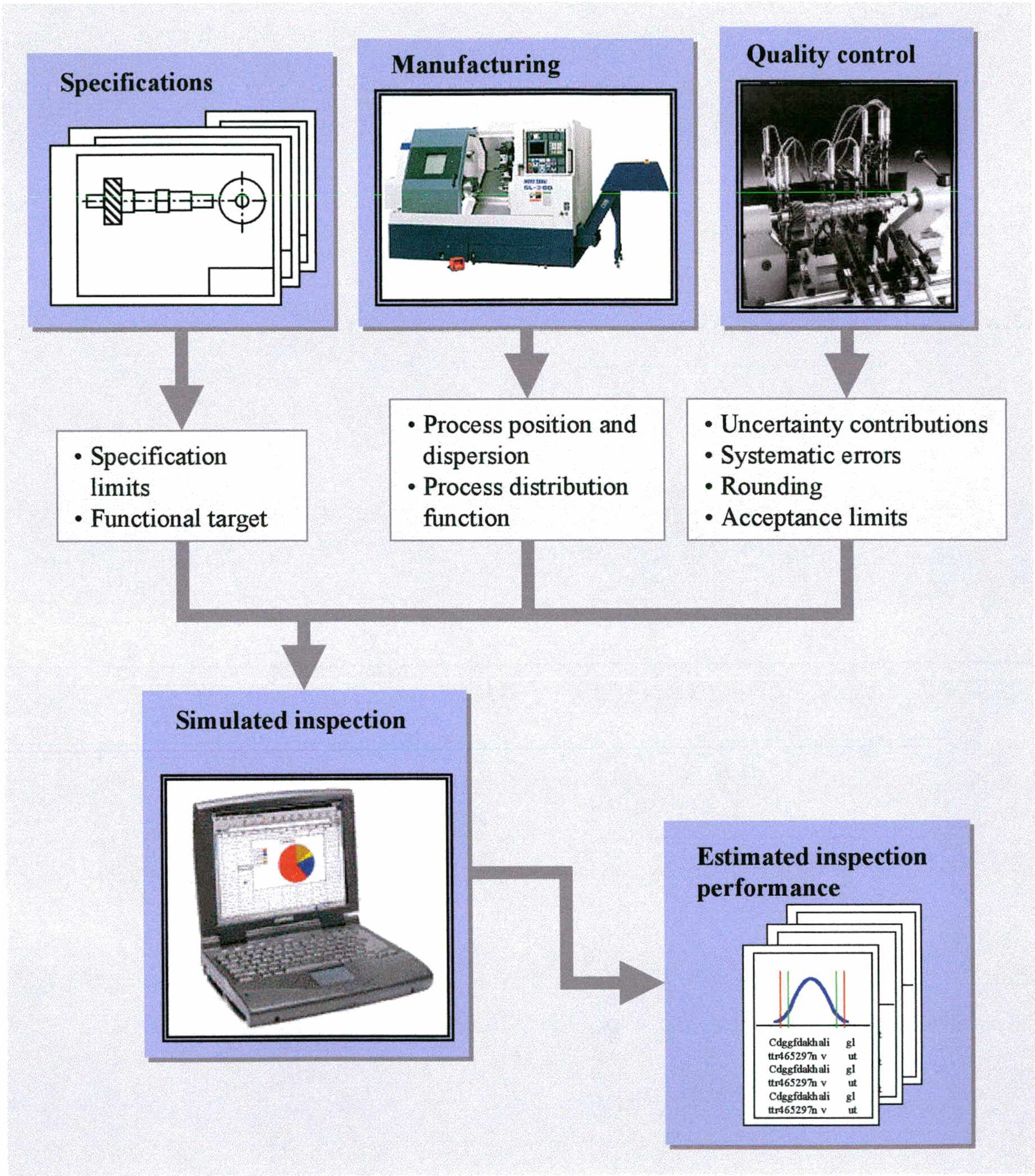

Figure 4.1: Data flow for the evaluation of inspection performance by simulation.

A simplified block diagram of the simulation algorithm is depicted in figure 4.2. It should be noted that the sequence of operations is similar to that of the experiment proposed in §3.3

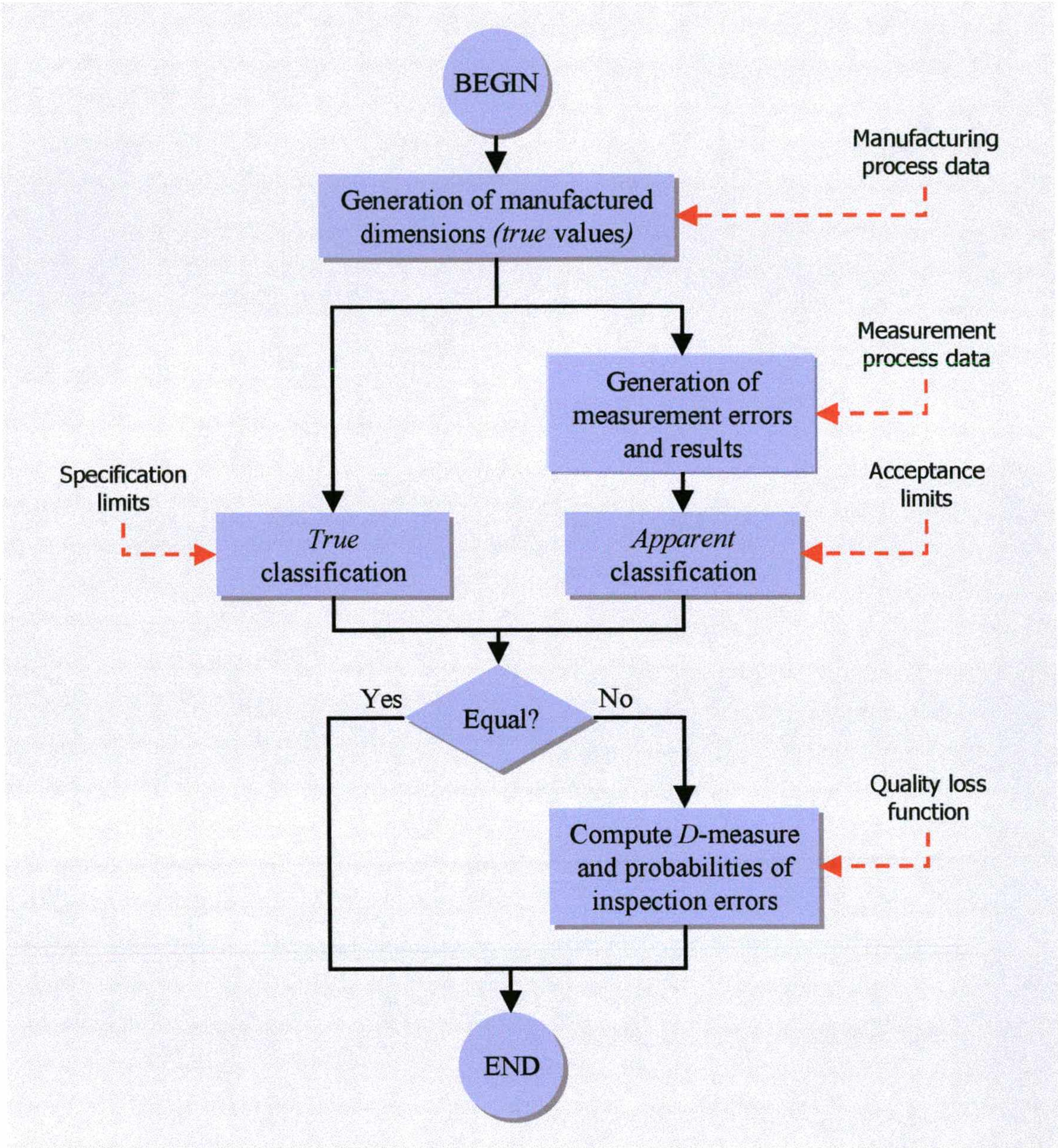

Figure 4.2: Simplified block diagram of simulated inspection.

The basic elements of the problem are:

- A set of rules to obtain the inputs to the algorithm from quantities available on the shop floor.
- Mathematical models of manufacturing and measurement processes.
- A set of rules to separate the simulated units into classes (*true* and *apparent* classification processes).

- A set of rules to identify inspection errors and to combine them in several measures of inspection performance.

These elements are described in the following sections.

4.1 Stochastic model of industrial measurement

A mathematical model of measurement process relates three basic quantities: the value of quality characteristic x, measurement error ε and measurement result y. As it has been already expressed in equation (30):

$$y = x + \varepsilon$$

These three quantities are of a random nature. In manufacturing and metrology, random variation can be reduced by proper design and construction of physical means (machine or instrument), by correct procedures and by close control of environmental quantities affecting repeatability /55, 56, 57/. The recognition of this possibility leads to the so-called *deterministic approach in manufacturing and metrology* /58/. However, random variability cannot be fully suppressed and its effects on conformity assessment should be considered by inspection system evaluation procedures. Then, it is possible to express equation (30) in terms of random variables (in this thesis random variables are noted uppercase):

$$Y = X + \mathrm{E} \tag{58}$$

It is widely accepted that total measurement error can be separated into its systematic and random components. Systematic errors are assumed to remain constant between measurement events within a given time interval. This makes them predictable and, to some extent, allows their correction. The value of some systematic error components can be determined by calibration of instrument with standards or master parts, operation that is performed in several points within the measurement range. The result of this operation is an error curve, which is formally defined only in the calibrated points and on a sample basis. This curve can be used, if convenient, to correct measurement results. However, correction is never complete, due to sampling and interpolation effects and other sources of uncertainty affecting calibration and correction processes. This results in a residual error that is systematic in nature and could be variable in the measuring range, with an unknown function. In addition, there are systematic error components that can not be quantified by calibration, e.g. long-term temperature variations, incorrect values of constants, etc. These unknown systematic errors, as well as the above-mentioned residual errors, are generally considered by type B contributions in uncertainty budgets built according to the GUM (see figure 4.3).

Random errors are variable from one measurement event to another, being not predictable. This is the case of the lack of repeatability, the effect of roughness and form deviations on length measurements, between-parts temperature variations, etc. These errors are taken into account by type A or B contributions in uncertainty budgets built according to the GUM (figure 4.3).

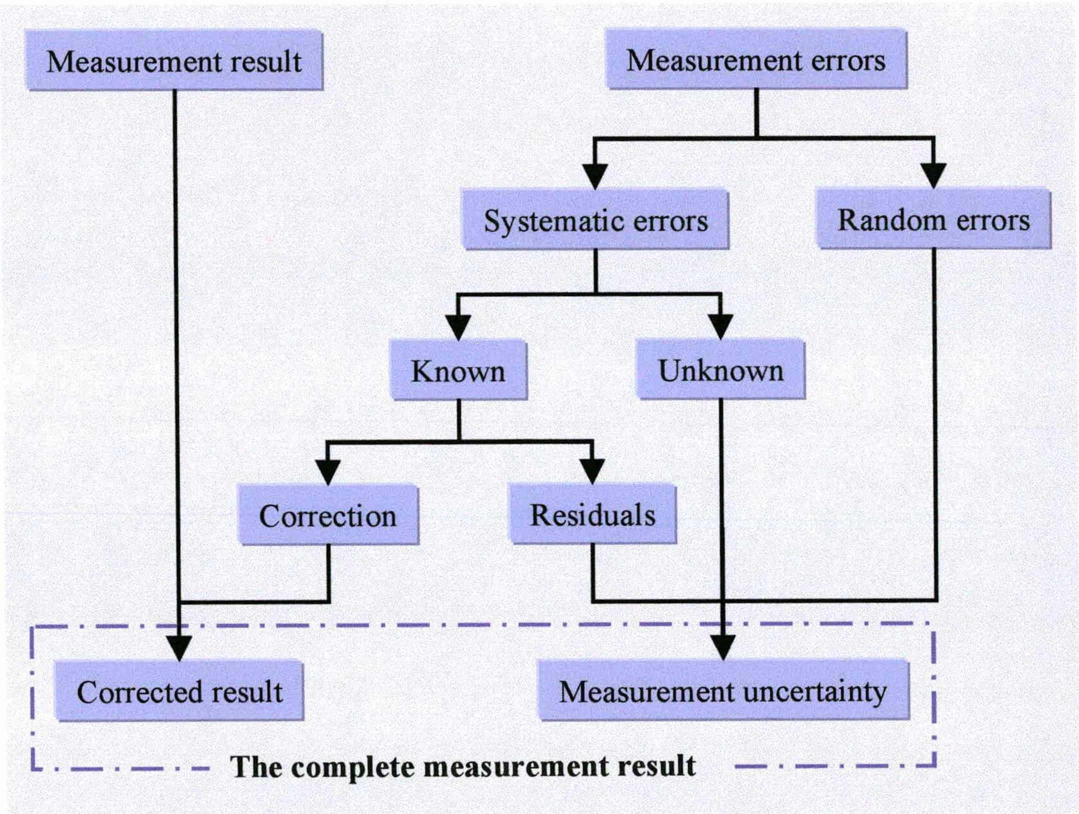

Figure 4.3: Contributions to measurement uncertainty (based on /17/).

Then, measurement error can be expressed as:

$$E = E_{ran} + E^{u}_{sys} + E^{k}_{sys} \tag{59}$$

where E_{ran}, E^{u}_{sys}, E^{k}_{sys} are variables representing random errors, unknown systematic errors and known (uncorrected) systematic errors respectively. It should be observed that systematic errors are treated here as random variables. Why? Because these errors are functionally dependent on the value of the quality characteristic, that is a random variable as well. This does not mean that they have a random behaviour in each position within the measuring range. However, events of systematic error produced when measuring a set of manufactured characteristics show a distribution. The shape of this distribution depends on the shape of the distribution of quality characteristic and on the shape of the systematic error curve.

In principle, the value of measurement errors can be represented by a continuous random variable. However, measurement results are always rounded to some extent. The mechanism by which rounding affects conformity assessment is different from that of systematic and random measurement errors /59/. Rounding produces a re-distribution of classification errors type I and type II in each acceptance or class limit. In consequence it will be treated here separately.

Based on the concepts above, the following mathematical model is proposed to simulate measurement of manufactured batches.

$$Y = \text{round}\left(X + \mathrm{E}_{ran} + \mathrm{E}_{sys}^{u} + \mathrm{E}_{sys}^{k}\right) \tag{60}$$

where round( ) is a rounding function. The variables X and E_{ran} are statistically independent, but E_{sys}^{u} and E_{sys}^{k} are functionally dependent on the variable X. The illustration of these dependencies through a covariance (or correlation) matrix in not useful, because dependencies are not linear.

In the following sections the stochastic models for these particular quantities are presented, along with the results of the validation tests performed.

4.1.1 Manufactured quality characteristics

The normal or gaussian PDF is generally adopted for the statistical characterisation of manufactured quality characteristics. This assumption is based on empirical, theoretical and operational reasons:

- Manufacturing practice shows that in many cases the distributions of manufactured quality characteristics do not depart considerably from normal.

- It could be considered that manufacturing deviations are produced by the combination of many error sources. So, regardless of the probability density of each source, the combined deviation becomes normally distributed as a consequence of the central limit theorem /60/.
- Normal PDF is defined by only two parameters: a mean μ_x and a variance σ_x^2.
- Normal statistics are simple and well known: it is easy to estimate the parameters of an hypothesised normal from the results of a small number of experiments.

This explains why most statistical process control tools rely on process normality /61/.

However, normal distribution should not be used as the only model to describe manufacturing process variability. It has been suggested that distributions of manufactured dimensions can vary from a rectangular to an approximately normal /62/, their limits are finite and they are not always symmetrical about their mean /63/. A typical example can be found in the inspection of units manufactured to fulfil form, position and orientation tolerances (e.g. circularity, cylindricity, perpendicularity, parallelism, run-out, etc.). The corresponding deviations present distributions that concentrate most events close to zero, which is the minimum possible value. Another objection comes from the analysis of the tails of normal PDF, which are particularly influent when dealing with conformity assessment. It has been suggested that the tails of the distributions of actual processes are not as predicted by the gaussian model /64/. In consequence, the evaluation of inspection performance based exclusively on the assumption of normality could fail to fit reality.

The beta PDF has been proposed by several authors as an alternative model for manufacturing processes /51, 63, 65/. The mathematical expression of the generalised beta depends on four parameters: upper and lower bounds of dispersion range (α and β respectively) and two exponents defining the distribution shape (λ_1 and λ_2). For a generic beta random variable T, it could be expressed as follows:

$$\begin{aligned} &\mathrm{f}(t)=\frac{1}{(\beta-\alpha)\cdot \mathrm{B}(\lambda_1,\lambda_2)}\cdot\left(\frac{t-\alpha}{\beta-\alpha}\right)^{(\lambda_1-1)}\cdot\left(1-\frac{t-\alpha}{\beta-\alpha}\right)^{(\lambda_2-1)} && \forall\{\alpha\le t\le\beta\} \\ &\mathrm{f}(t)=0 && \forall\{t<\alpha \;\; or \;\; t>\beta\} \end{aligned} \tag{61}$$

where t are events of a random variable T and $\mathrm{B}(\lambda_1;\lambda_2)$ is the value of the complete beta function, which can be defined in terms of the gamma function as:

$$\mathrm{B}(\lambda_1,\lambda_2)=\frac{\Gamma(\lambda_1)\cdot\Gamma(\lambda_2)}{\Gamma(\lambda_1+\lambda_2)} \tag{62}$$

The beta PDF is the model that allows representing distributions with the broadest spectrum of skewness and kurtosis /66/. Symmetric distributions are obtained by exponents of equal value and can vary from U-shaped to normal. Asymmetric distributions require the exponents to be different and can vary from J-shaped to a skewed-bell shape.

The complexity of the beta PDF is not a problem for its use in the simulation algorithm. The generation of beta-distributed random numbers is simple, though it is not as time-efficient as the generation of normal deviates (e.g. see /67/ and /68/). Fitting a beta PDF to sample data is also non-problematic (the procedure adopted in this thesis is described in §6.1.2). The only problem that restricts the application of the beta PDF in practical cases is the lack of statistics to infer confidence intervals for population parameters from the estimated sample parameters. Because of this, the use of the beta model should be limited to those cases for which large samples are available. Then, the beta random variable representing the value of manufactured quality characteristic is:

$$X\sim \mathrm{B}(\alpha;\beta;\lambda_1;\lambda_1) \tag{63}$$

In spite of the objections above, the normal distribution is also used in this thesis to describe the behaviour of manufactured quality characteristics:

$$X\sim \mathrm{N}(\mu_x;\sigma_x^2) \tag{64}$$

Because of its simplicity, it will be applied in the study of the behaviour of the D-measure (chapter 5). It will be also offered as an option in the industrial evaluation package, to be used when the sample size is not enough to apply the beta PDF.

The statistical properties of the normal PDF can be found in any probability textbook. A comprehensive treatment of the properties of beta distribution is available in /69/.

4.1.2 Random measurement errors

Measurement and manufacturing are not essentially different processes. In consequence, the considerations in §4.1.1 are still valid when discussing about random measurement errors.

The blind use of the normal PDF has also been criticised in metrology /70/. In spite of that, it can be applied to represent the variability of measurement results when the number of error contributions is large. In this thesis the normal model will be used to represent random errors in the evaluation of the performance of the *D*-measure (chapter 5). It will be also offered as an option in the industrial evaluation package. In this case the mean or expected value μ_{ran} is equal to zero. Thus:

$$\mathrm{E}_{ran} \sim \mathrm{N}\left(0;\sigma_{ran}^2\right) \tag{65}$$

Other probability density functions are proposed to be used in particular circumstances. *A priori* evaluation of the inspection performance could require a more conservative model of random errors. For those cases, the rectangular PDF could be the right choice:

$$\mathrm{E}_{ran} \sim \mathrm{R}(-\xi;+\xi) \tag{66}$$

where $[-\xi,+\xi]$ is the interval around zero within which the error is expected to lie. Less conservative analysis can be performed using the triangular PDF in the same interval:

$$\mathrm{E}_{ran} \sim \mathrm{Tr}(-\xi;+\xi) \tag{67}$$

The statistical properties of these density functions can be found in any basic probability textbook or in the GUM /23/.

4.1.3 Unknown and residual systematic measurement errors

A proper model of systematic errors must consider the functional relationship between the variable E_{sys}^u and the value of the measurand *X*. Nevertheless, E_{sys}^u is expected to represent systematic contributions to uncertainty: its functional relationship with *X* is unknown. It can only be stated that, for each point in the measuring range of the instrument, the value of systematic error will be within an interval $[h-,h+]$. Usually, in uncertainty budgets built according to the GUM, it is assumed that the probability density of any value of systematic error within the interval $[h-, h+]$ is constant. In addition, it is generally considered that the interval is symmetrically placed at zero, being so $h- = h+ = h$.

It is worth remembering that, whatever the functional relationship between X and E^{u}_{sys}, it remains approximately constant for all units measured within a given time interval. In consequence, the model in this thesis represents unknown and residual systematic errors as curves determined by random parameters. During simulation, each curve is used to compute the values of E^{u}_{sys} to be replaced in equation (60) for all the events of X. Afterwards, the operation is repeated for several curves to evaluate the effect of the lack of knowledge on the relationship between X and E^{u}_{sys} (this procedure will be described in more detail in §4.2).

The model has been tuned to fulfil the following set of requirements:

- If X is normally distributed, the curves have to be continuous and defined in $(\mu_x - 6\cdot\sigma_x) \leq x \leq (\mu_x + 6\cdot\sigma_x)$. If X is beta-distributed, the curves must be defined in the interval $\alpha \leq x \leq \beta$.
- For a given limit value of systematic error h, the total amplitude of the generated curves can be any, within a range $\pm h$.
- The curves should reflect the superposition of a constant component and a calibration-like variation.
- For each dimension x in the definition interval, the distribution of systematic error values obtained by the repeated generation of curves should be approximately trapezoidal, with amplitude $\pm h$.

For any particular error curve, the events of systematic error are calculated as a function of the value of the measurand and the amplitude of the interval within which the curve must lie:

$$\varepsilon^{u}_{sys} = \Phi(x;h) = \frac{h}{3}\cdot(k_c + \Psi(x)) \tag{68}$$

where k_c is the constant part of the error and $\Psi(x)$ is the calibration-like variation. For normally distributed manufacturing processes, the function $\Psi(x)$ is defined by:

$$\Psi(x) = 2 \cdot k_1 \cdot \sin\left\{2 \cdot \pi \cdot \left[\frac{k_3}{12} \cdot \left(6 + \frac{x - \mu_x}{\sigma_x}\right) + k_4\right]\right\} + k_2 \cdot \sin\left\{2 \cdot \pi \cdot \left[\frac{1 + k_5}{12} \cdot \left(6 + \frac{x - \mu_x}{\sigma_x}\right) + k_6\right]\right\} \tag{69}$$

where k_i, i=1...6, are parameters that modify the shape and relative amplitude of the curve. Thus, different functions $\Psi(x)$ can be obtained changing the values of the parameters k_i as events of rectangularly distributed random variables $K_i \sim \mathrm{R}(0;1)$.

The constant part of the error k_c is also generated as an event of a uniformly distributed random variable. The limits of this variable are computed from the excess of amplitude available within the interval $\pm h$:

$$\begin{aligned} &\frac{\max[\Psi(x)] + \min[\Psi(x)]}{2} \le 0 \qquad \Rightarrow \qquad K_c \sim \mathrm{R}(-\{3 + \min[\Psi(x)]\};0) \\ &\frac{\max[\Psi(x)] + \min[\Psi(x)]}{2} > 0 \qquad \Rightarrow \qquad K_c \sim \mathrm{R}(0;\{3 - \max[\Psi(x)]\}) \end{aligned} \tag{70}$$

where $\min[\Psi(x)]$ and $\max[\Psi(x)]$ are the minimum and maximum values of the calibration-like variation in the definition interval $\mu_x - 6 \cdot \sigma_x \le x \le \mu_x + 6 \cdot \sigma_x$.

In figure 4.4 a sample of eight error curves generated by the model described in equations (68)-(70) can be observed. The variable X has been changed into a non-dimensional variable $Z \sim \mathrm{N}(0;1)$, being the relationship with the variables of the problem defined by $z = (x - \mu_x)/\sigma_x$. Thus, the definition interval is $-6 \le z \le 6$. Systematic error is reported also in terms of the non-dimensional variable W, where $w = \varepsilon_{sys}^u / h$.

It can be observed that all the error curves in figure 4.4 are consistent with the statement:

$$-1 \le w \le 1; \qquad \forall z \mid \quad -6 \le z \le 6 \tag{71}$$

or, in terms of the variables of the problem:

$$-h \le \varepsilon_{sys}^u \le h; \qquad \forall x \mid \quad \mu_x - 6 \cdot \sigma_x \le x \le \mu_x + 6 \cdot \sigma_x \tag{72}$$

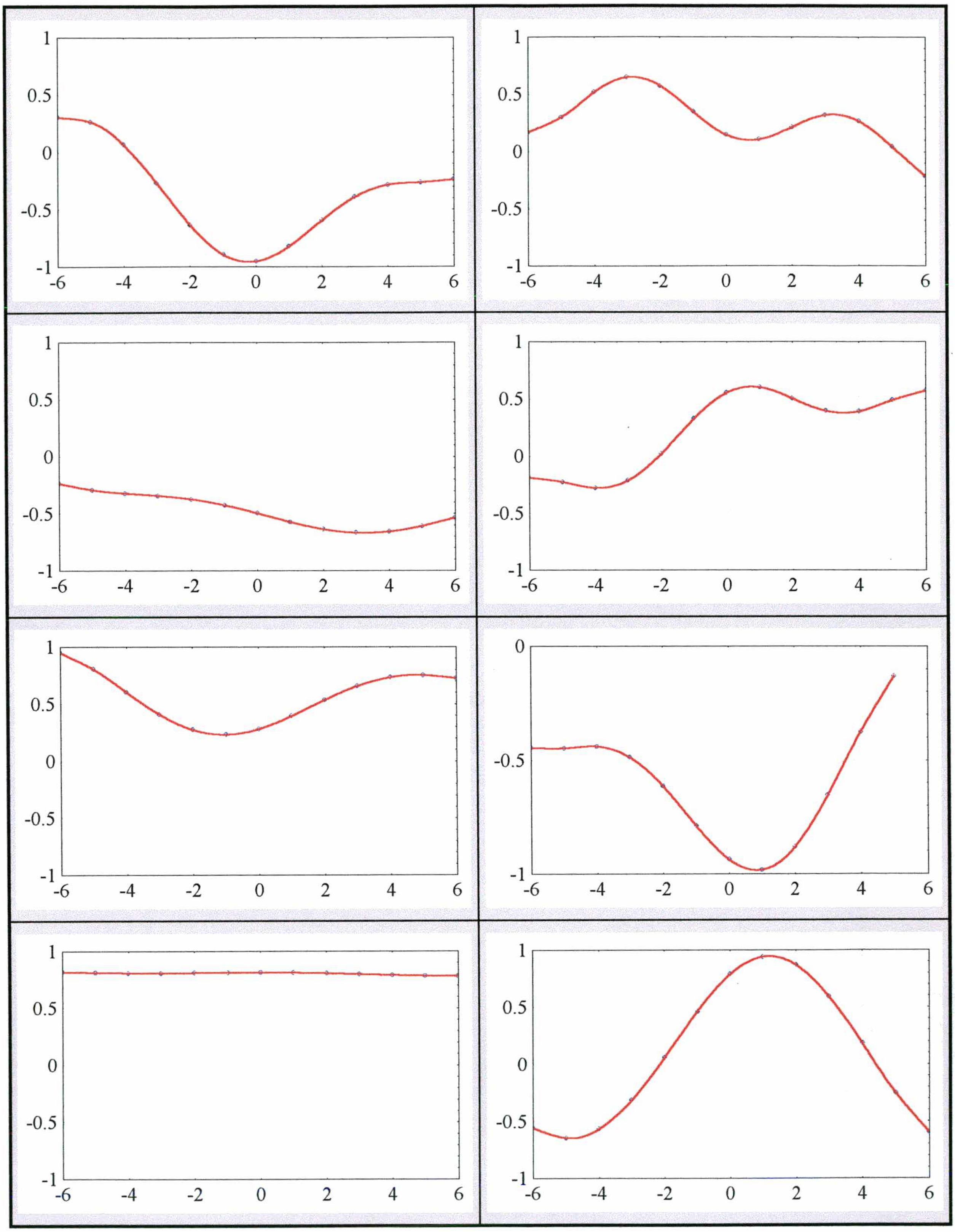

Figure 4.4: Sample of eight error curves generated applying the model in equations (68)-(70)

To complete the validation of the model against the premises, 1000 curves have been sampled in different values of the non-dimensional quality characteristic z. The distribution of non-dimensional errors in an arbitrary value of z is plotted in figure 4.5.

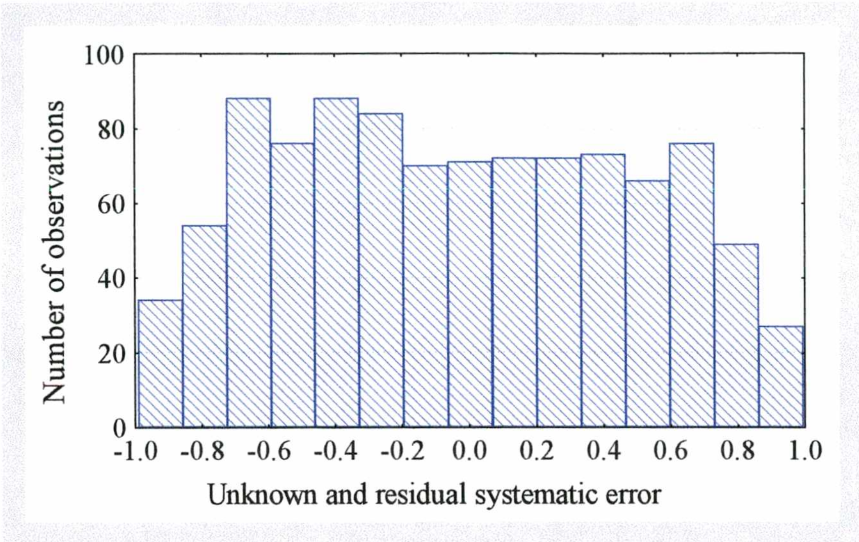

Figure 4.5: Distribution of systematic errors in $z = 0$ $(x = \mu_x)$.

The shape of this distribution remains approximately constant within the definition interval of the quality characteristic. Its standard deviation is:

$$\sigma_w\Big|_{z=\text{constant}} \cong \frac{1}{2} \tag{73}$$

or, in terms of the variables of the problem:

$$\sigma_{\varepsilon_{sys}^u}\Big|_{x=\text{constant}} \cong \frac{h}{2} \tag{74}$$

The stochastic model in equations (68)-(70) can be easily modified for beta-distributed processes. Equations however are not shown in this document.

The proposed model fulfils the requirements set above. It could be criticised from different points of view, but particularly concerning the shape of the generated error curves. It could be asked: are these curves valid representations of unknown and residual systematic errors? The answer to this question is not known, because there is no empirical evidence on how these components behave. Nevertheless, it should be recognised that the curves generated by the model are physically feasible. In addition, any set of curves is consistent with a statement on systematic contributions to uncertainty according to the GUM. Thus, the model can be considered sufficient for the purpose of this thesis.

4.1.4 Non-corrected (known) systematic measurement errors

The GUM prescribes that all recognised systematic effects must be corrected. However, the nature of systematic errors in industrial metrology makes it difficult to achieve this ideal situation. Some relevant facts are:

- Systematic errors that are variable in the measuring range are difficult to correct unless some kind of automatic data acquisition and processing system is available (this is not always the case).
- Systematic errors do not remain necessarily constant during long production runs /71/.
- The lack of repeatability and reproducibility and/or a poor resolution of the indicating device hide the *true* value of systematic error. The uncertainty introduced in the correction by these contributions can be greater than the error itself.

Because of the reasons above, the complete correction of systematic effects is not frequent in industrial metrology. At best a partial correction is made, which consists in zeroing the errors in a point of the measuring range using a standard or master part. This operation is performed periodically, to control variations in time.

It is known that systematic measurement errors affect more the quality of outgoing product than random errors. From this viewpoint, it seems necessary to consider their effects in the assessment of inspection performance. Two different cases arise regarding the moment in which this assessment is made. During *a priori* assessment, systematic errors can not be known for each point in the measuring range. The only information available comes from instrument specifications, normally in the form of a maximum error. This situation is similar to that of unknown and residual systematic errors. Thus, the model in §4.1.3 should be used.

Once the inspection system is available, the error curve can be estimated by experimentation (calibration). It could be the case that a complete correction were difficult to achieve and a partial correction were preferred (e.g. zeroing the error in a point of the measuring range). Then, it could be interesting to simulate the effect of residual error on the inspection performance. To make it possible, a continuous function has to be fitted to the error values. This function must be able to represent the behaviour of the actual measurement system in the calibration points but also in any other point within the measuring range.

In this thesis, a polynomial model is proposed to fit and interpolate the values of systematic error obtained by calibration. Because of empirical reasons, the degree of the polynomial has been restricted to six:

$$\varepsilon_{sys}^{k} = \sum_{i=0}^{m} c_i \cdot x^i \; ; \qquad m \leq 6 \tag{75}$$

The coefficients of the polynomial are computed by means of the least squares method (OLS). The degree of the polynomial is defined within the iterative procedure depicted in figure 4.6.

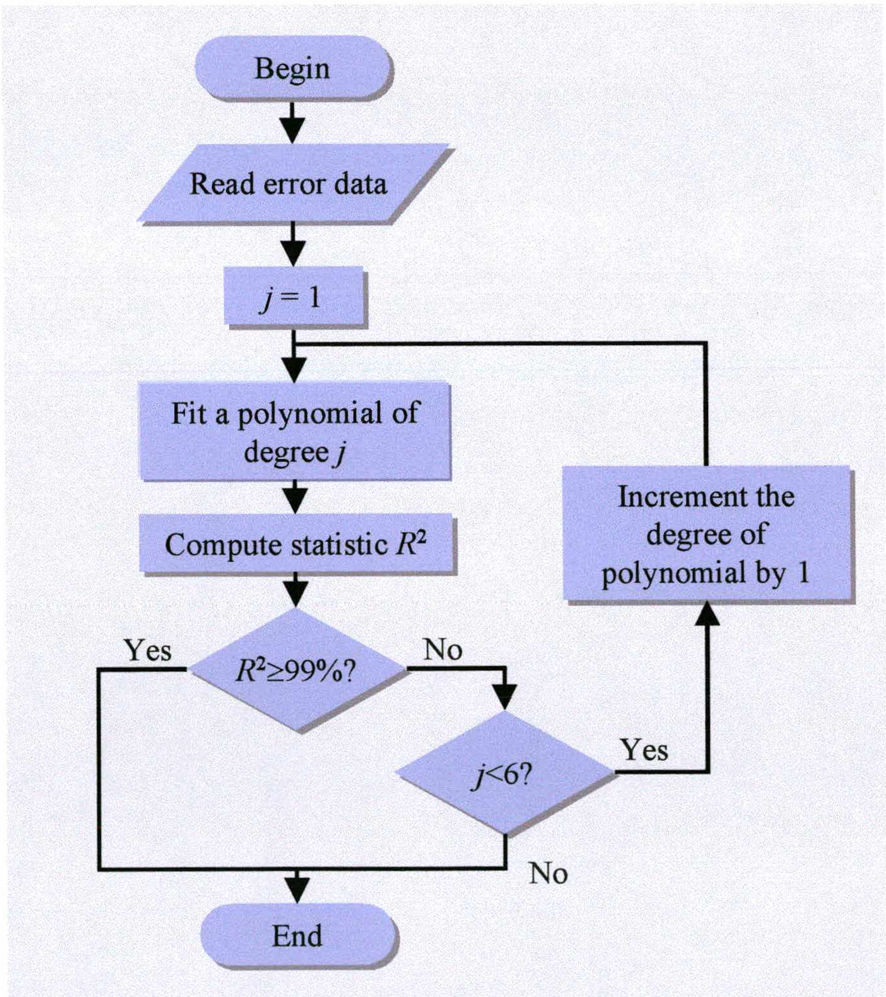

Figure 4.6: Fitting procedure for known systematic errors

The control parameter is the goodness of fit, indicated by the coefficient R^2. In this case, R^2 is the percentage of the sum of squares of systematic error about its mean explained by the polynomial curve. In equations:

$$R^2 = \frac{\sum_{i=1}^{m}\left(\hat{\varepsilon}_{sys_i}^{k} - \bar{\varepsilon}_{sys}^{k}\right)^2}{\sum_{i=1}^{m}\left(\varepsilon_{sys_i}^{k} - \bar{\varepsilon}_{sys}^{k}\right)^2} \cdot 100\% \tag{76}$$

The procedure begins with a polynomial of degree 1 –straight line- and stops when the desired quality of fit has been obtained ($R^2 \geq 99\%$) or when the degree of polynomial is 6, whatever is satisfied first. To test the fitting procedure, 22 calibration certificates have been processed. The selected instruments were micrometers, dial and digital gages and lever-type gages (e.g. Pupitast). The fitting results of several calibration certificates are illustrated in figure 4.7.

The obtained values of R^2 are within the interval $54.18 \leq R^2 \leq 99.18$, being the arithmetic mean $\overline{R^2} = 86.21\%$. Despite $\overline{R^2}$ is acceptable, the lowest value of the dispersion interval indicate poor fitting results. It can be shown that all cases with a low value of R^2 correspond to calibration curves that present drastic variations between adjacent calibration points (see cases 2 and 8, figure 4.7). Case 2 shows a predominant effect of resolution that hides the value of the systematic error. Case 8 shows an anomalous behaviour of the instrument in the lower third of measuring range. Nevertheless, in all studied cases the curves lie within the calibration uncertainty band. Then, the fitted polynomial provides feasible error values.

The goodness of fit can be improved by increasing the degree of the polynomial. However, goodness of fit is not the only criterion to measure the success of a regression operation. A better fit would satisfy one of purposes of regression: replacing a set of values by a formula, to be used only for the values of the experiment. But the relationship could fail to satisfy the second major purpose of regression, which is interpolation /72/. Indeed, in the measured points the problem of choosing an adequate regression relationship is of statistical nature, but between points it is largely empirical. If the degree of the polynomial were increased beyond 6, high curvatures could occur between experimental values to follow abrupt variations like those in cases 2 and 8. These variations should not be followed, but filtered.

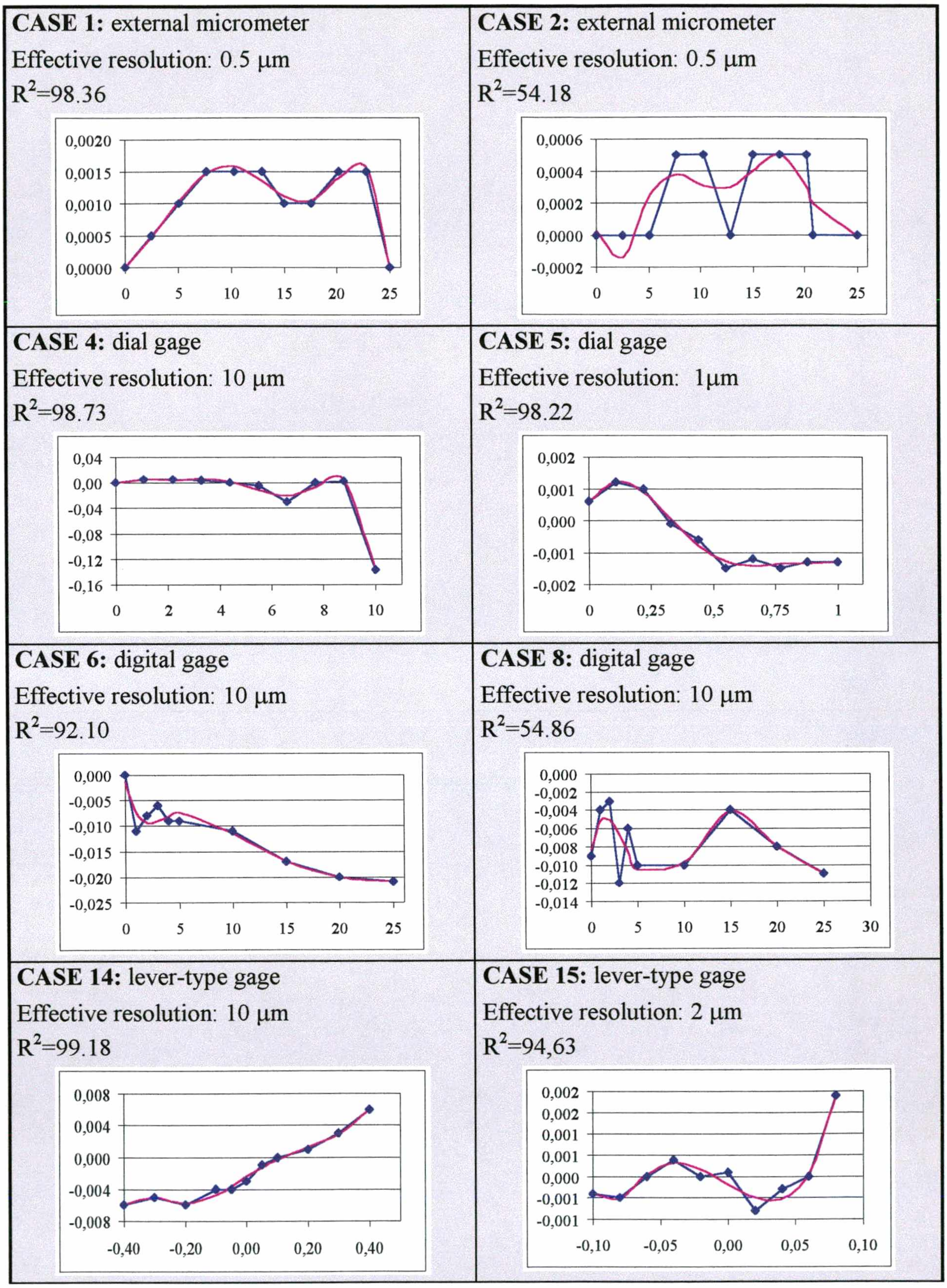

Figure 4.7: Results of the polynomial fitting in several calibrated instruments

In consequence, the model in equation (75) will be used in this thesis together with the procedure in figure 4.6. From the metrological viewpoint, the use of this procedure is not recommended unless it can be shown that the error is stable in time. Otherwise the inspection performance could change, making the prediction not useful. In case of doubt, it is better to apply the model for unknown and residual systematic errors (§4.1.3).

4.1.5 Rounding

Resolution is defined as the smallest difference between indications of a displaying (or recording) device that can be meaningfully distinguished /10/.

In case of digital displays, the value of resolution is equal to the change in indication when the least significant digit changes by one step. The measurement error introduced when reading a digital indicating device is of a systematic nature. The value of the measurand, already affected by other error components, is presented to the inspector rounded to the nearest integer number of resolutions. This can be described by a saw-teeth error pattern that is stable in time and can be considered constant in measuring range (see figure 4.8).

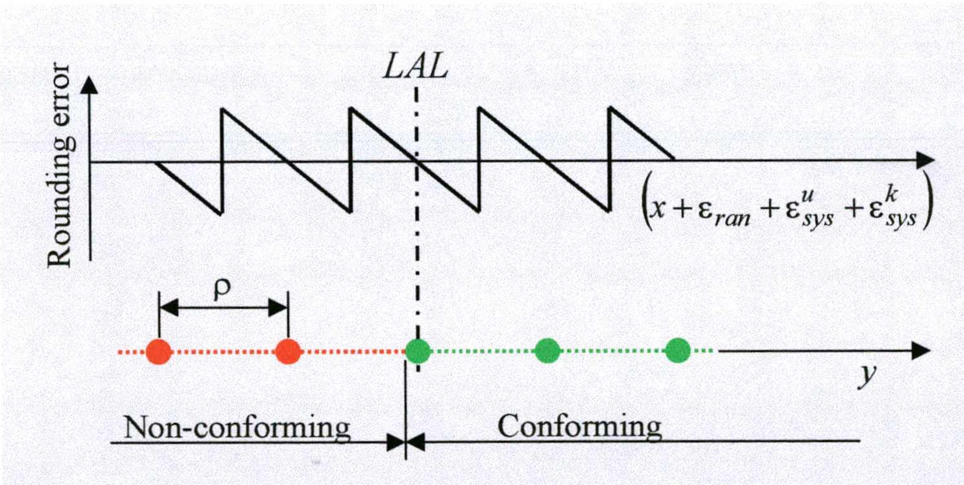

Figure 4.8: Effect of resolution in digital indicating devices.

This type of error is itself a source of misclassification. According to standard practices and recommendations, units producing measurement results equal to acceptance limits should be considered conforming /28/. For example, units producing measurement results equal to the lower acceptance limit could correspond to values of the input signal within the following interval (see figure 4.8):

$$LAL - \rho/2 < \left(x + \varepsilon_{ran} + \varepsilon_{sys}^{u} + \varepsilon_{sys}^{k}\right) < LAL + \rho/2 \qquad (77)$$

where *LAL* is the lower acceptance limit and ρ is the resolution. It should be noted that half of this interval is in the non-conforming zone, thereby increasing the probability of accepting non-conforming units.

The resolution of analogue displays is determined by the scale increment, but also by the level of eye interpolation that can be achieved under a particular reading condition. If the condition is favourable the minimum scale increment can be divided meaningfully in up to ten intervals. In that case, the resolution will be one-tenth of the scale increment. There is no model available of the total error produced when reading analogue displays. This error can be separated into three main components: misreading the scale, parallax and incorrect interpolation of fractional divisions. Regarding interpolation, it has been observed that in the proximity of dial marks inspectors can read the *true* position of pointer with almost no bias and small uncertainty /73/. This condition is similar to that produced in conformity assessment applications, where the position of pointer has to be compared with the marks in the dial or scale representing acceptance limits. Thus, interpolation error seems to be a minor source of misclassification. Unfortunately, other error components like parallax are not easy to quantify and to model. Their effect on inspection performance is not known, though there are no reason to consider it negligible. In addition, analogue displays create favourable conditions for other operator-induced error known as flinching, which is the tendency to falsify the results of borderline products /3/. In general, during inspection by variables flinching produces that non-conforming units in the neighbourhood of the acceptance limit are reported as being conforming. This is equivalent to a non-formal, non-controlled, enlargement of the acceptance interval, with unknown consequences on product quality.

After this brief discussion on reading errors, it can be concluded that modelling the metrological behaviour of analogue displays requires a particular investigation that is out of the scope of this thesis. In consequence, it has been decided to use a simple model to simulate the effects of rounding:

$$y = \rho \cdot \mathrm{int}\left(\frac{x + \varepsilon_{ran} + \varepsilon_{sys}^{u} + \varepsilon_{sys}^{k}}{\rho} + \frac{1}{2} \right) \tag{78}$$

where int() is a function that returns the integer of argument between brackets. This model corresponds to the already discussed behaviour of digital displays. Then, it will fit for automatic inspection equipment. In that case the value of ρ is the resolution at which the comparison among measurement result and acceptance limits is performed. The model can also be used to simulate the behaviour of analogue displays, though it is not particularly appropriate for this task. In this case, the value of ρ should be replaced by the effective resolution and not by the minimum scale increment. Anyway, this restriction does not seem to be very conflictive because today inspection systems use mostly digital indicating devices.

4.2 Simulated inspection

In figure 4.9, the procedure to evaluate inspection performance is depicted. The first step, not shown in the flowchart, is to define the type of inspection operation and enter the required data. Afterwards, the algorithm generates a set of *true* values of the manufactured quality characteristic. These values are compared one at a time with the specification limits, providing attributes that describe the *true* classification status of each quality characteristic.

The first step towards the generation of measurement results is to consider, if appropriate, the effect of non-corrected systematic errors. To do this a polynomial is fitted to the calibration results using the procedure in §4.1.4. The obtained function is used to compute events of systematic error for the generated *true* values. The effect of random and systematic contributions to measurement uncertainty is evaluated by repeated simulation (*r* times). Each time the loop is executed, a different curve is used to compute events of unknown and residual systematic error. These curves are defined using the procedure in §4.1.3. Events of random errors are changed also for each loop, but the parameters of their distribution remain constant. Finally, measured values are obtained summing the errors with the corresponding *true* values and rounding them to the nearest integer number of resolutions. This way, a different set of measurement results is obtained each time the loop is executed. Nevertheless, all these sets are composed by measurement results that are consistent with the available knowledge.

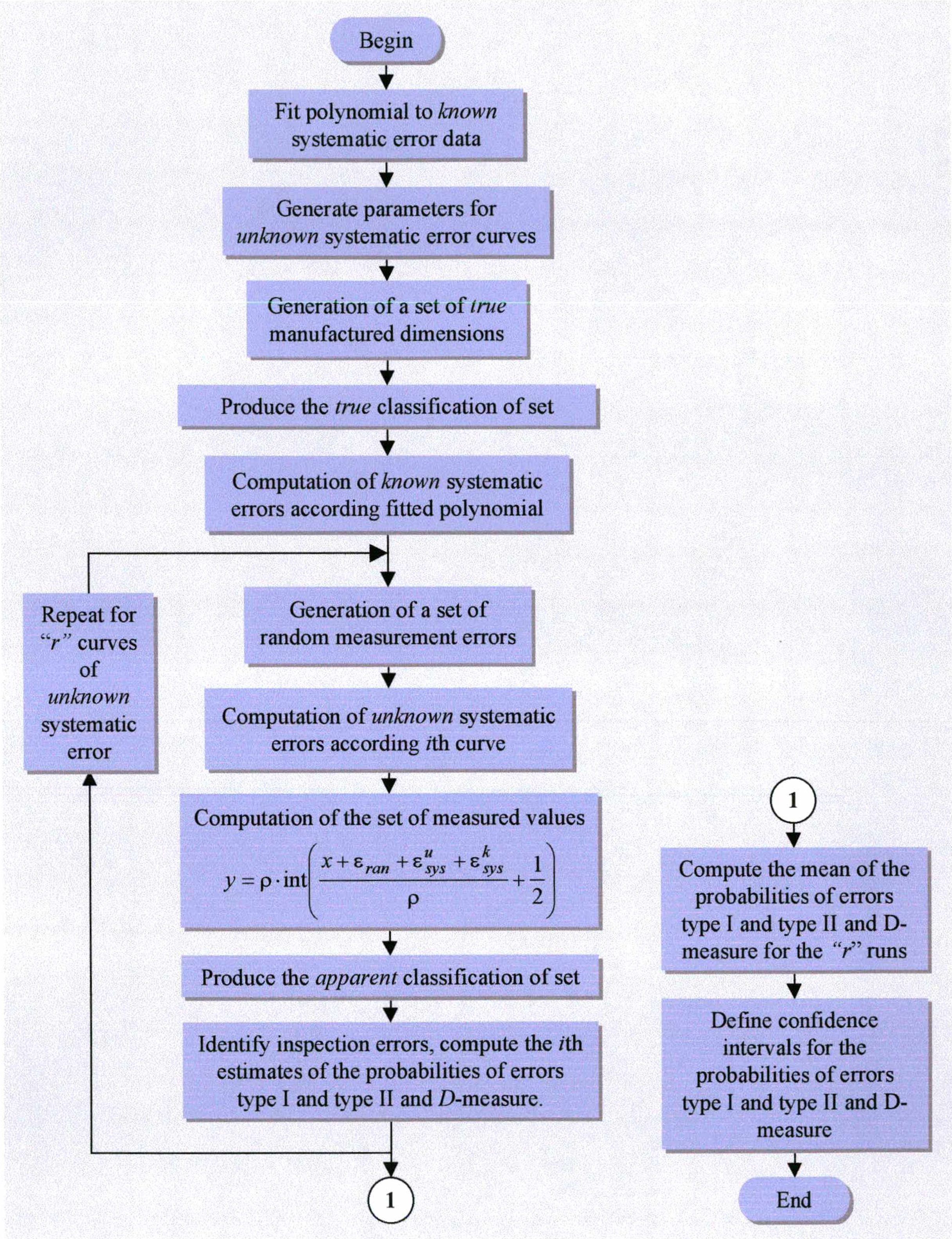

Figure 4.9: Flowchart for the evaluation of inspection performance by simulation.

Events of measurement results are classified with respect to acceptance limits, generating a set of attributes that describe the *apparent* classification status of the quality characteristics. Afterwards, inspection errors are identified comparing these attributes with those obtained in the *true* classification operation. Table 4.1 provides the decision rules to perform the *apparent* classification and identify inspection errors considering the effect of rounding (see §2.2 for the continuous variable approach). Note that the separation in types I and II has been maintained, though that it is not required for the estimation of the D-measure.

Inspection error	**Definition**	**Conditions for occurrence**	
		$AL_i \in$ class at left	$AL_i \in$ class at right
Type I^{SL_i}	A unit belonging to the class at left of SL_i is reported in the class at right	$x < SL_i \wedge y > AL_i$	$x < SL_i \wedge y \geq AL_i$
Type II^{SL_i}	A unit belonging to the class at right of SL_i is reported in the class at left	$x > SL_i \wedge y \leq AL_i$	$x > SL_i \wedge y < AL_i$

Table 4.1: Definition and conditions for the occurrence of errors type I and type II.

For the sake of simplicity, the class for which the specification limit SL_i is the upper limit is called "class at left of SL_i". Similarly, the class for which SL_i is the lower limit is called "class at right of SL_i". It should be noted that conditions for *apparent* classification and error identification change depending upon whether the acceptance limit belongs to the class at right or at left. These decision rules are valid regardless of the type of inspection operation. The concepts of class and limit among adjacent classes are the same in the inspection of one- or two-sided tolerances or in dimensional classification: they do not depend on whether the class is conforming or non-conforming.

The sets of variables and attributes generated in each loop are used to estimate the values of the D-measure and probabilities of error type I and type II. This process is schematically depicted in figure 4.10, for an inspection operation involving m specification limits.

The estimated values of D for each loop are noted by $\hat{D}j$, $j = 1..r$. The equations for their computation can be found in §3.1.1 to §3.1.5, depending on the inspection case and the relative position of the target.

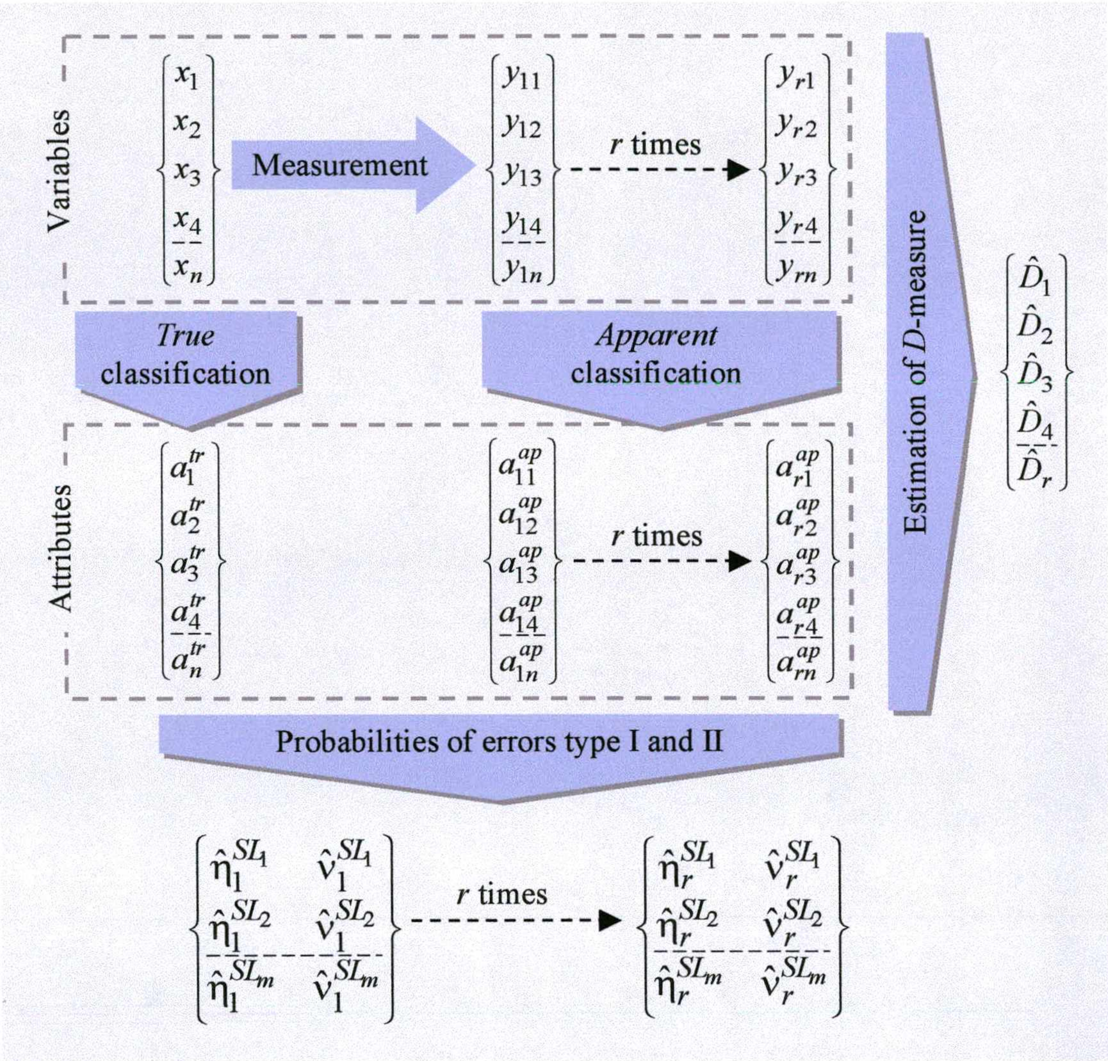

Figure 4.10: The estimation of the D-measure and the probabilities of error type I and type II.

It should be remembered that the r loops are not repetitions of the same process. The curve representing unknown and residual systematic errors is changed from one loop to another. Thus, the set $\{\hat{D}j\}$ informs on the inspection performance of a family of measurement systems that satisfy the same conditions regarding measurement uncertainty. The statistical properties of this set can be used to create a 95% confidence interval for the expected value of D-measure:

$$\mathrm{E}\{D\} = \overline{D} \pm t_{r-1;\,97.5\%} \cdot \sqrt{\frac{s_D^2}{r}} \tag{79}$$

where $\overline{D}$ is the sample mean:

$$\overline{D} = \frac{1}{r} \cdot \sum_{j=1}^{r} \hat{D}_j \tag{80}$$

and s_D^2 is the sample variance:

$$s_D^2 = \frac{1}{r-1} \cdot \sum_{j=1}^{r} \left(\hat{D}_j - \overline{D}\right)^2 \tag{81}$$

and $t_{r-1;97.5\%}$ is the value of the t-student variable for $(r-1)$ degrees of freedom and 95% confidence level (two-sided interval, 5% out). In all simulations associated with this thesis, the number of loops is $r = 50$, so that $t_{r-1;97.5\%} = 1.96$.

The confidence interval in (79) should not be confused with the uncertainty in the value of the D-measure. All the values in $\{\hat{D}j\}$ are consistent with the available knowledge on measurement conditions, expressed by measurement uncertainty. Then, it is necessary to create an interval that includes the values of the D-measure that could be expected in a given inspection operation.

This is not an easy task, because the probability distribution of $\{\hat{D}j\}$ is not normal. As the number of loops is big enough ($r = 50$), it has been decided to approximate the dispersion interval using the difference between extreme values in the sample:

$$\min\left(\hat{D}_j\right) \leq D \leq \max\left(\hat{D}_j\right) \quad ; \quad \forall\, j = 1 \ldots r \tag{82}$$

where $\min\left(\hat{D}_j\right)$ and $\max\left(\hat{D}_j\right)$ are respectively the minimum and maximum values of the D-measure obtained in the 50 loops.

The probabilities of inspection errors type I and type II can be treated in a similar way. The following equation defines the estimated probability of error type I for each specification limit and loop:

$$\hat{\eta}_j^{SL_i} = \frac{1}{n} \cdot [\text{counted number of errors type I}]_j^{SL_i} \tag{83}$$

where n is the size of the simulated batch. The sub-index j identifies the loop number, $j = 1..r$. The supra-index SL_i identifies the specification limit under analysis, $i = 1..m$ (m is the number of specification limits). The 95% confidence interval for the expected value of the probability of error type I can be defined by:

$$E\{\eta^{SL_i}\} = \overline{\eta}^{SL_i} \pm t_{r-1;\,97.5\%} \cdot \frac{s^2_{\eta^{SL_i}}}{\sqrt{r}} \tag{84}$$

where $\overline{\eta}^{SL_i}$ is the sample mean:

$$\overline{\eta}^{SL_i} = \frac{1}{r} \cdot \sum_{j=1}^{r} \hat{\eta}_j^{SL_i} \tag{85}$$

and $s^2_{\eta^{SL_i}}$ is the sample variance:

$$s^2_{\eta^{SL_i}} = \frac{1}{r-1} \cdot \sum_{j=1}^{r} \left(\hat{\eta}_j^{SL_i} - \overline{\eta}^{SL_i}\right)^2 \tag{86}$$

Finally, the dispersion interval is defined by:

$$\min\left(\hat{\eta}_j^{SL_i}\right) \le \eta_j^{SL_i} \le \max\left(\hat{\eta}_j^{SL_i}\right) \quad ; \quad \forall\, j = 1..r \tag{87}$$

Similar equations can be derived for the probability of error type II:

$$\hat{\nu}_j^{SL_i} = \frac{1}{n} \cdot [\text{counted number of errors type II}]_j^{SL_i} \tag{88}$$

$$E\{\nu^{SL_i}\} = \overline{\nu}^{SL_i} \pm t_{r-1;\,97.5\%} \cdot \frac{s^2_{\nu^{SL_i}}}{\sqrt{r}} \tag{89}$$

$$\overline{\nu}^{SL_i} = \frac{1}{r} \cdot \sum_{j=1}^{r} \hat{\nu}_j^{SL_i} \tag{90}$$

$$s^2_{\nu^{SL_i}} = \frac{1}{r-1} \cdot \sum_{j=1}^{r} \left(\hat{\nu}_j^{SL_i} - \overline{\nu}^{SL_i}\right)^2 \tag{91}$$

$$\min\left(\hat{\nu}_j^{SL_i}\right) \le \nu_j^{SL_i} \le \max\left(\hat{\nu}_j^{SL_i}\right) \quad ; \quad \forall\, j = 1..r \tag{92}$$

The values of ϕ and θ can be estimated for each loop using equations (24) and (25):

$$\hat{\phi}_j = \frac{1}{\hat{p}} \cdot \left[\hat{\nu}_j^{LSL} + \hat{\eta}_j^{USL} \right] \tag{93}$$

$$\hat{\theta}_j = \frac{1}{\hat{q}} \cdot \left[\hat{\eta}_j^{LSL} + \hat{\nu}_j^{USL} \right] \tag{94}$$

where:

$$\hat{p} = \frac{\text{number of conforming units}}{\text{total number of units}} \tag{95}$$

$$\hat{q} = 1 - \hat{p} \tag{96}$$

The quantities in equations (93)–(96) can be used to estimate the measures of inspection performance in §2.1.

5 BEHAVIOUR AND APPLICATION OF THE *D*-MEASURE

One of the aims of this thesis is to determine the relationship between the quality of measurement results and the deterioration of product quality due to non-ideal inspection. The former is associated with the value of measurement uncertainty and the latter to the incremental quality loss that can be attributed to inspection errors. In this chapter, results of a large number of simulated inspection tasks are presented and discussed. Conclusions are drawn about the validity of capability statements based on the *D*-measure, measurement uncertainty and other indices of inspection performance.

5.1 Simulation domain and procedure

The simulation procedure described in chapter 4 has been applied to evaluate the behaviour of the *D*-measure under a broad set of conditions (1000 parameter combinations). The study has been made only for 100% inspection of two-sided specifications, which is the most common case in production metrology. It has been assumed that the manufacturing target is in the middle of tolerance interval. Thus, the nominal-the-best QLF has been used to compute the values of the *D*-measure.

The manufacturing process distribution has been considered normal, with its mean centred on the functional target. This makes the problem symmetric: the same behaviour should be expected in both specification limits. The concept of process capability has been applied to relate the standard deviation of process with the tolerance interval [61]:

$$Cp = \frac{T}{6 \cdot \sigma_x} \tag{97}$$

Process capability determines the quality loss and affects the classification performance. Considering that 100% inspection is used mainly when $Cp < 1.33$, the following capability levels have been established for the study: 0.67 (p=95.44%), 0.83 (p=98.76%) and 1.00 (p=99.73%).

It is assumed that known systematic errors have been corrected. Only random (σ_{ran}) and unknown systematic errors (h) remain, that combine with measuring device resolution (ρ) to determine the value of measurement uncertainty:

$$U_{95} \cong 2 \cdot \sqrt{\sigma_{ran}^2 + 0.25 \cdot h^2 + \frac{\rho^2}{12}} \tag{98}$$

To simplify the study, non-dimensional variables have been created dividing all the problem variables by the tolerance and bringing the manufacturing target to zero. The definition of these variables and the values that determine the simulation domain are detailed in table 5.1.

Parameter	Definition	Simulation domain/values
Tolerance	------	$T^* = 1.0$
Manufacturing target	------	$m^* = 0.0$
Specification limits	$LSL^* = \frac{LSL - m}{T}$ $USL^* = \frac{USL - m}{T}$	$LSL^* = -0.5$ $USL^* = 0.5$
Mean of manufacturing process	$\mu_x^* = \frac{\mu_x - m}{T}$	$\mu_x^* = 0.0$
Standard deviation of manufacturing process	$\sigma_x^* = \frac{\sigma_x}{T} = \frac{1}{6 \cdot Cp}$	$\sigma_x^* = \frac{1}{4};\ \frac{1}{5};\ \frac{1}{6}$
Resolution of measurement system	$\rho^* = \frac{\rho}{T}$	$\rho^* = 0.001;\ 0.005;\ 0.01;\ 0.05;\ 0.1$
Measurement uncertainty	$U_{95}^* = \frac{2}{T} \cdot \sqrt{\sigma_{ran}^2 + 0.25 \cdot h^2 + \frac{\rho^2}{12}}$	$U_{95}^* \sim \mathrm{R}\left(\frac{\rho^*}{\sqrt{3}}; 0.2\right)$
Interval for unknown and residual measurement errors	$h^* = \frac{h}{T}$	$h^* \sim \mathrm{R}\left(0; \sqrt{U_{95}^{*\,2} - \frac{\rho^{*2}}{3}}\right)$
Standard deviation of random measurement error	$\sigma_{ran}^* = \frac{\sigma_{ran}}{T}$	$\sigma_{ran}^* = \frac{1}{2} \cdot \sqrt{U_{95}^{*\,2} - \frac{\rho^{*2}}{3} - h^{*2}}$
Limit displacements	$Ld^* = \frac{LAL - LSL}{T} = \frac{USL - UAL}{T}$	$Ld^* = \rho \cdot \mathrm{int}\left[\frac{1}{\rho} \cdot \mathrm{R}\left(0; \frac{3 \cdot U_{95}^*}{2}\right)\right]$

Table 5.1: Definition of simulation variables and description of simulation domain.

Standard deviation of manufacturing process and resolution of measuring device are chosen at random among the values listed in table 5.1. After that, uncertainty of measurement is defined as an event of a rectangular distribution. The lower limit is determined by the effect of rounding, the upper limit by 0.2 ($U_{95} \leq T/5$). The values of σ_{ran} and h are defined at random, within the remaining uncertainty. A scatter plot of the 1000 simulated combinations of random and unknown systematic errors is depicted in figure 5.1.

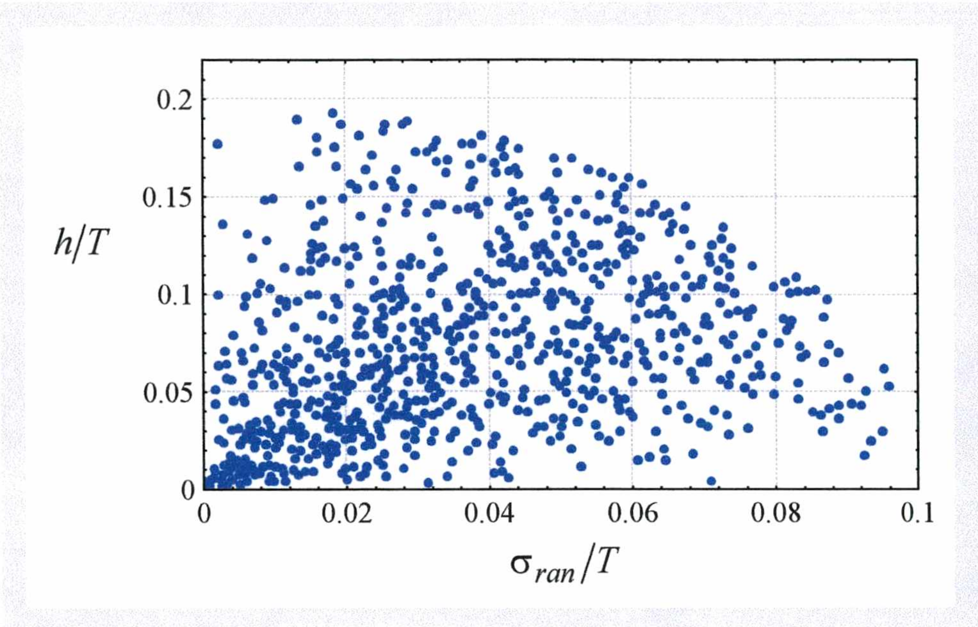

Figure 5.1: Thousand simulated values of σ_{ran}/T and h/T.

It should be noted that an ellipse determines the bound of this domain. The error combinations on this bound correspond to the smallest possible resolution. Higher resolution values result in error combinations that are closer to the origin of the plot.

The same value of displacement has been applied in both specification limits (see table 5.1). This value is defined at random between zero and 1.5 times measurement uncertainty. A rounding function has been applied to assure that the acceptance limits are placed on an integer number of resolutions. In the analysed case, units whose measured values are equal to the acceptance limits are assumed to fulfil the specification.

For the 1000 combinations of parameters, values of the D-measure, λ^r, ϕ and θ (mean and extremes in 50 simulation loops) have been computed. The results are presented and analysed in the following sections.

5.2 The relationship between *D* and measurement uncertainty

Figure 5.2 shows the mean values of the D-measure [%] obtained for measurement systems with different uncertainties (U_{95}/T has been used instead of U_{95}). Each point in the plot corresponds to a particular set of parameters in the simulation domain described in table 5.1.

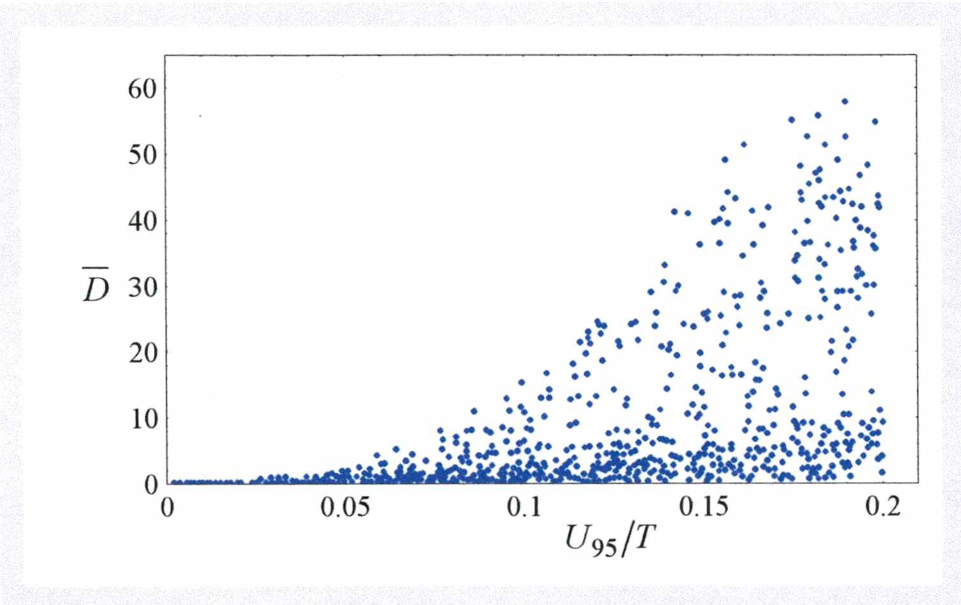

Figure 5.2: The relationship between $\overline{D}$ [%] and the uncertainty per tolerance ratio (1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

It can be observed that different values of $\overline{D}$ have been obtained for a given value of uncertainty per tolerance ratio. The dispersion grows with the value of U_{95}/T, mainly because of the effect of limit displacements. Indeed, the highest values of $\overline{D}$ are caused by massive rejection of conforming units that occurs when the displacements are close to $1.5 \cdot U_{95}/T$. To verify this statement 235 cases without displacement of limits have been separated from the main data set. These cases are plotted in figure 5.3.

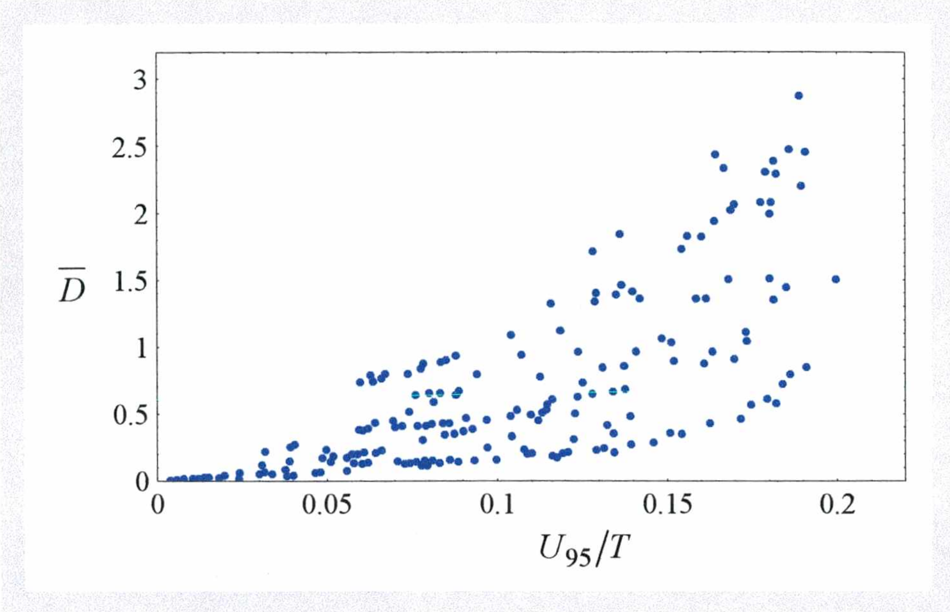

Figure 5.3: The relationship between $\overline{D}$ [%] and the uncertainty per tolerance ratio (235 cases without limit displacement).

The scatter of $\overline{D}$ values is still present, but in a completely different scale. Even if the measurement system has an uncertainty $U_{95} = 0.2 \cdot T$ the fraction of quality loss that can be attributed to inspection errors is less than 3%, that is a negligible amount.

The plot in figure 5.3 reveals another important phenomenon. Observing carefully the position of points, three parabolic patterns can be detected. It can be shown that points that are close to the upper bound of the dispersion region correspond to cases in which manufacturing processes operate with low capability ($Cp = 0.67$). As the process capability grows, the mean values of the D-measure decrease. Thus, the lowest parabolic pattern corresponds to manufacturing processes with $Cp = 1.00$. It should be remembered that the higher the process capability, the lower the total quality loss with ideal inspection λ^i. This way, figures 5.2 and 5.3 mix cases with three different values of total quality loss with ideal inspection. This can be clearly perceived if the total quality loss with non-ideal inspection λ^r [%] is plotted against U_{95}/T (see figure 5.4). When no limit displacements are applied, the influence of measurement system on the total quality loss is almost negligible if compared with the influence of process capability.

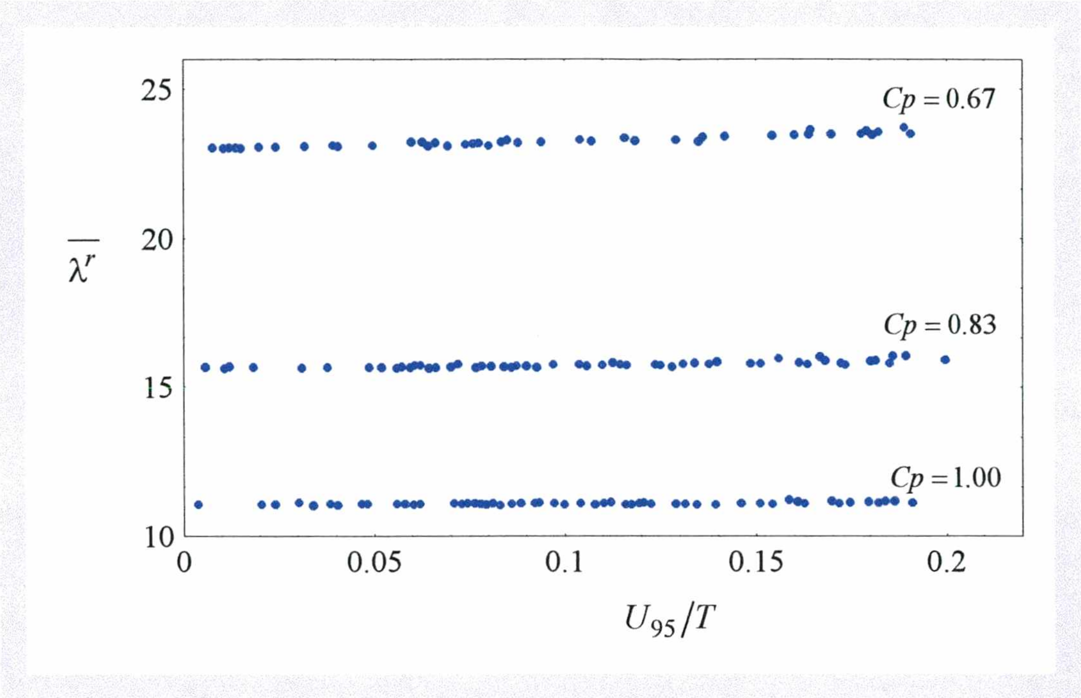

Figure 5.4: Relationship between the total quality loss and the uncertainty per tolerance ratio (235 cases without limit displacements).

The picture changes if limit displacements are used to preserve the quality of inspected units (see figure 5.5). If $U_{95} \leq T/10$, the patterns remain separated from each other: the process capability still dominates. This is reasonable, because small uncertainties imply small limit displacements (the condition was $USL - UAL \leq 1.5 \cdot U_{95}$). On the contrary, when $U_{95} > T/10$ the influence of limit displacements becomes relevant, even surpassing the effect of manufacturing process capability. This phenomenon is evidenced on the right side of figure 5.5. The ordered patterns vanish, leading to an almost random cloud of points.

It should be noted that the limit between both regions ($U_{95} = T/10$) is associated with the values of *Cp* used in the simulation. If a smaller increment between *Cp* values were used, the limit between the ordered- and random-behaviour regions would displace to smaller values of U_{95}/T. Eventually, this limit could disappear if *Cp* were treated as a random variable. In spite of that, it can be affirmed that measurement system-related decisions could define the amount of total quality loss, that is to say manufacturing economy, when limit displacements are used.

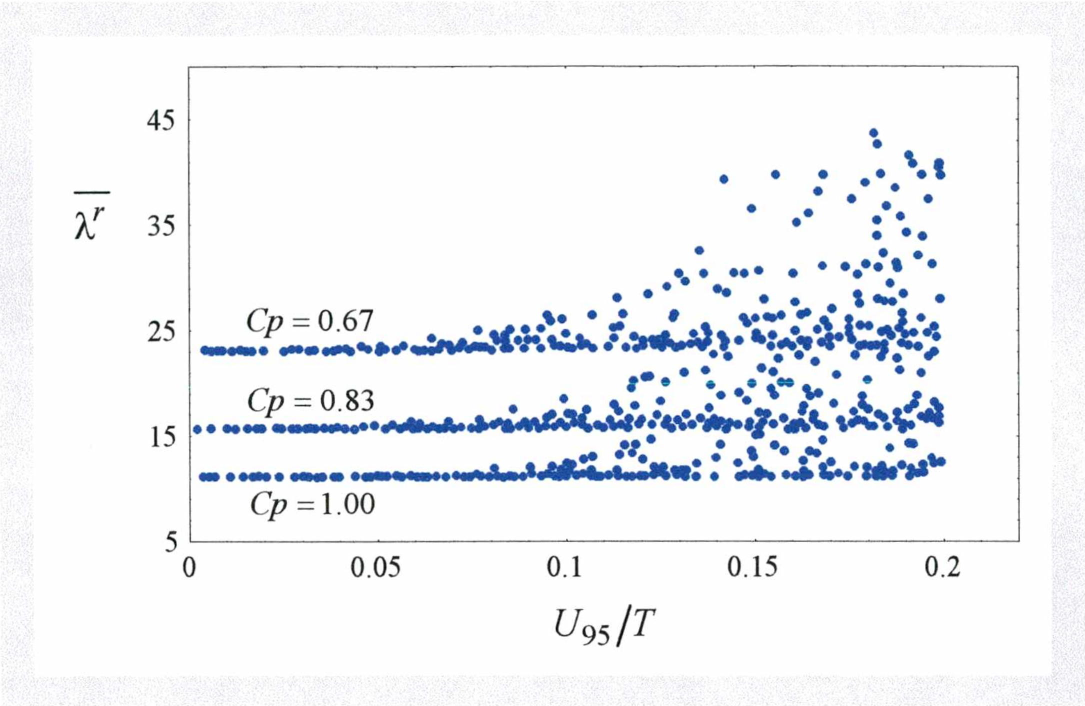

Figure 5.5: Relationship between the total quality loss and the uncertainty per tolerance ratio (1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

Up to this point, only arithmetic mean values of D and λ^r have been used. However, simulation results show that both quality measures scatter along the 50 loops of a simulation run. The first cause of this dispersion is the finite sampling of continuous random variables. A set of 10000 events of random measurement error is generated within each loop. So, even in absence of other contributions to uncertainty, the values of the D-measure will vary for each simulation loop. This variation is expected to be small, because the size of sample is large. The generation of *true* values of manufactured units is also affected by sampling. However, the effect does not appear as a scatter in the D-measure values, because only one set of quality characteristics is used for the 50 loops (see figure 4.9). In this case the error introduced by sampling is unknown, though it can be considered small because of the large sample size.

The second cause of scatter in the values of D is the lack of knowledge of systematic errors. It is known that systematic errors affect more the inspection performance than random errors of similar size. Thus, the random change in the error curve used to simulate unknown and residual systematic errors (see §4.1.3) is expected to be a major contributor to the scatter of D.

After this brief discussion, it seems reasonable to study the behaviour of the D-measure for different combinations of σ_{ran} and h. The range of values of the D-measure obtained in the 50 simulation loops has been used as a measure of the scatter:

$$\Delta = \max(D_j) - \min(D_j)\ ; \quad \forall j = 1 \ldots 50 \tag{99}$$

The behaviour of Δ has been studied for 235 cases without limit displacements. The results are shown in figure 5.6, plotted in a 2D domain defined by σ_{ran} and h and fitted with a polynomial of 2nd degree in both independent variables.

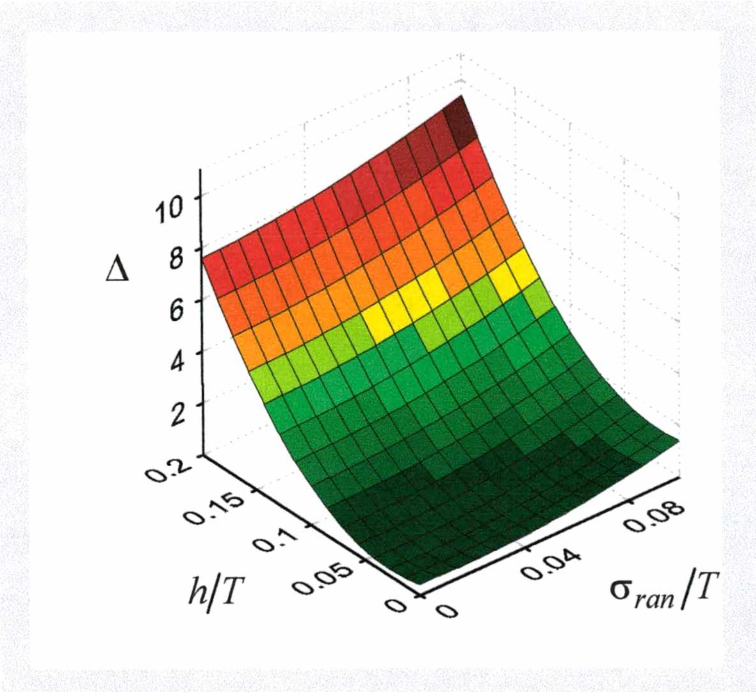

Figure 5.6: Effect of random and systematic contributions to uncertainty on the scatter of the D-measure (235 cases without limit displacements).

As expected, random errors have negligible effect on the variation of the D-measure. Thus, if the lack of repeatability is the major contribution to measurement uncertainty $\overline{D}$ will estimate the *true* inspection performance with little uncertainty. On the contrary, the dispersion in the values of D grows drastically when systematic contributions dominate the uncertainty budget. In this case $\overline{D}$ is a less precise estimator of the *true* inspection performance: it is associated with a higher uncertainty. The same behaviour has been found in the analysis of 1000 cases with limit displacements, though the values of Δ are higher than those in figure 5.6 (not shown).

It is important to interpret the results in figure 5.6 properly. The 235 cases plotted in the figure above correspond to different inspection conditions. Not only uncertainty contributions vary from case to case but also the manufacturing process capability and measuring device resolution. This produces a relevant dispersion of values of Δ with respect to the fitted surface. Then, the plot in figure 5.6 should not be used to estimate Δ for a particular measuring condition defined by $(\sigma_{ran};h)$. It is only a means to show that the lack of knowledge of systematic errors lead to inspection performances that are rather uncertain. In that case, the metrologist should define whether it is appropriate to use the value of $\overline{D}$ or the maximum value of $\hat{D}_j$ obtained in the 50 simulation loops.

5.3 Relationship between *D* and other measures of inspection performance

The aim of this section is to compare the values of several measures of inspection performance with the value of D that can be found under the same conditions. The study has been performed on the complete data set generated for §5.2, that is to say 1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$. For each case the values of $\bar{\phi}$ (mean probability of rejecting conforming units) and $\bar{\theta}$ (mean probability of accepting non-conforming units), as well as the estimates of p (fraction conforming) and q (fraction non-conforming), have been computed. Then, the equations in table 2.1 have been used to estimate the values of twelve measures of inspection performance. The results can be observed in figures 5.7 and 5.8. In each plot the value of the studied measure is represented on the horizontal axis, while the values of $\overline{D}$ [%] are on the vertical one.

It is not necessary to discuss each measure separately. The most important common characteristic is that none of them is completely correlated with the D-measure: there exists always some scatter. This scatter can not be attributed to sampling effects, because the compared measures have been calculated within the same simulation run. Therefore, the D-measure defines a scale of inspection performance that is not consistent with the scales defined by measures in table 2.1.

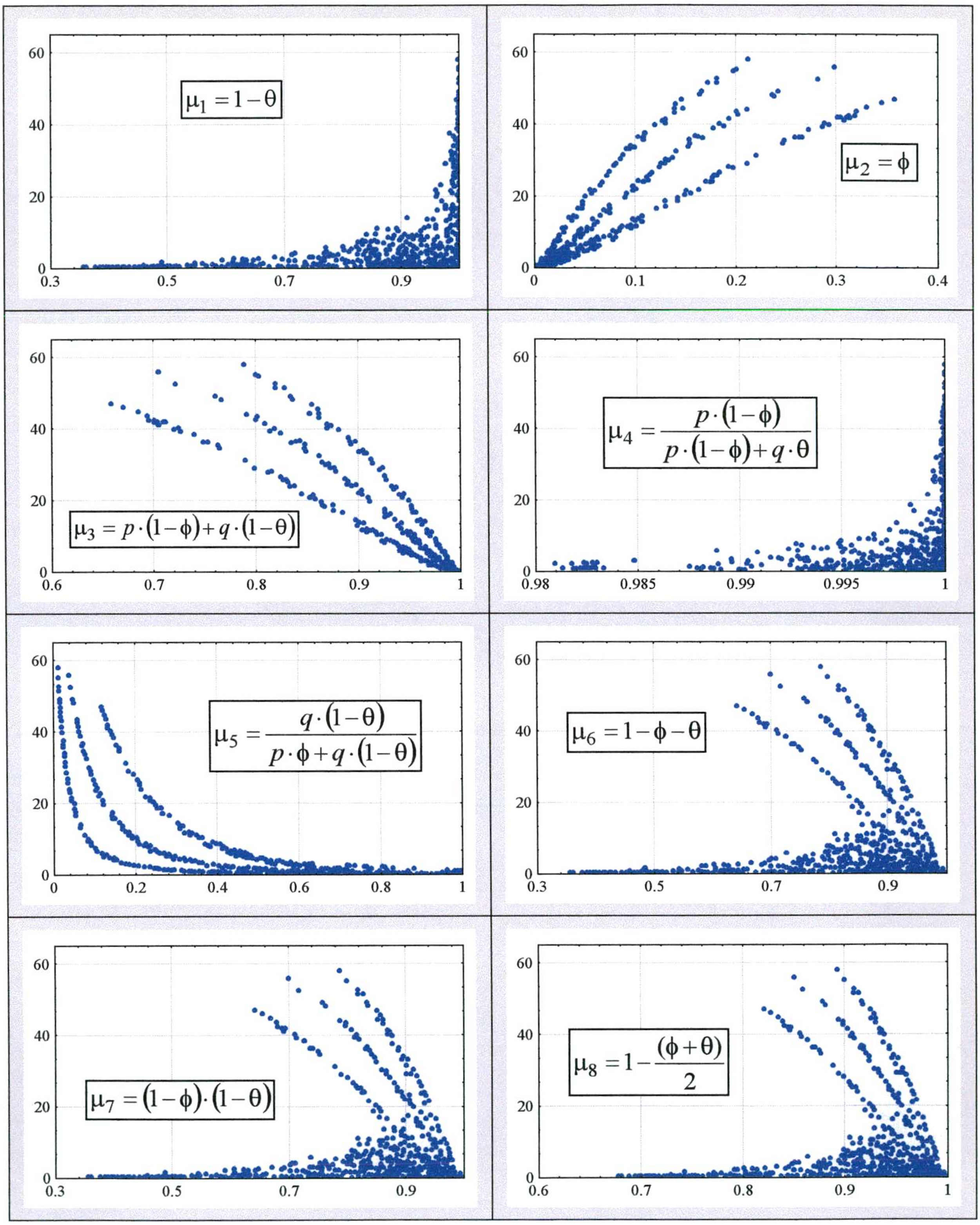

Figure 5.7: Relationship between D [%] and measures of inspection performance μ_1 to μ_8 (1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

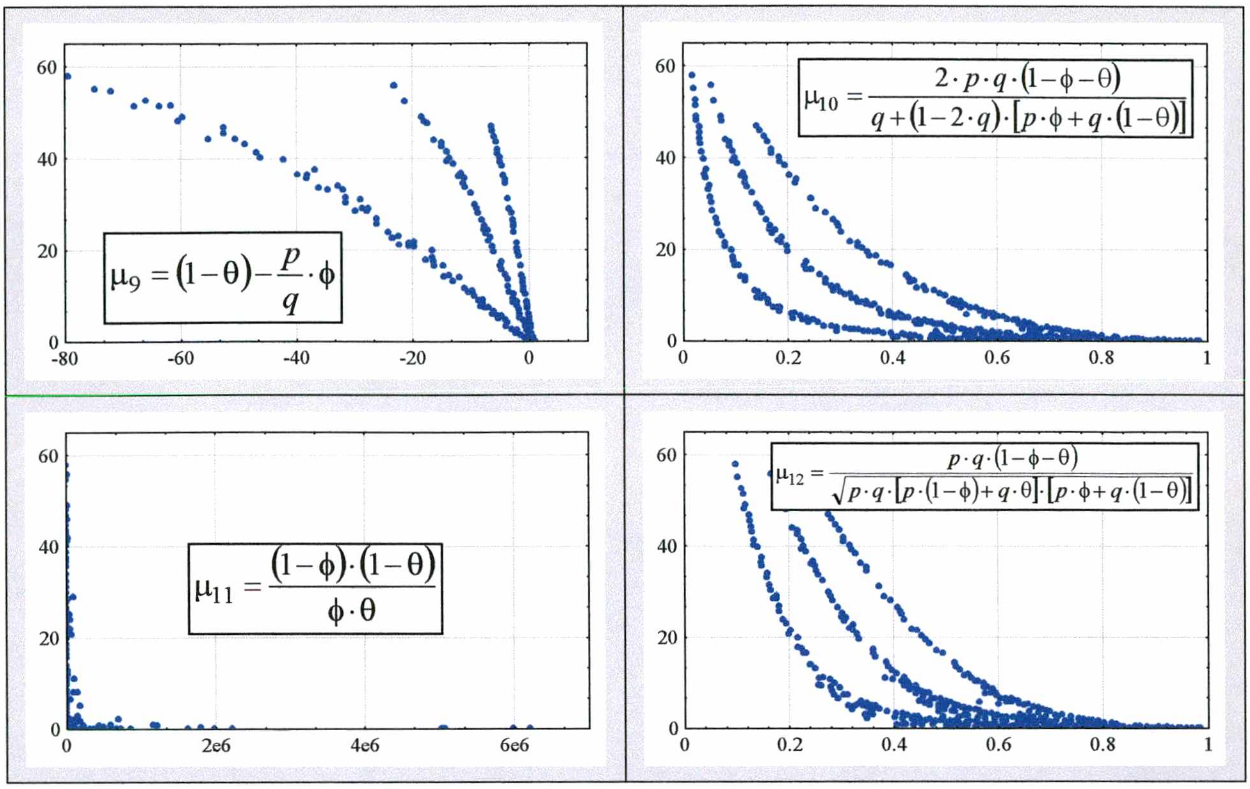

Figure 5.8: Relationship between D [%] and measures of inspection performance μ_9 to μ_{12} (1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

A remarkable fact is that the best correlation with $\overline{D}$ is provided by those measures dominated by the value of ϕ, probability of classifying a conforming unit as non-conforming. These measures are μ_2, μ_3, μ_5, μ_9, μ_{10} and μ_{12}.

The plot of μ_2 (figure 5.7) shows that the value of $\overline{D}$ grows with $\overline{\phi}$. Three clear patterns can be detected, corresponding to manufacturing processes that operate with different capabilities. The upper pattern corresponds to $Cp = 1.00$, the lower one to $Cp = 0.67$. It should be remembered that these patterns exist because only three values of Cp have been used for the simulation. If Cp were randomised, a cloud of points would replace the patterns. On the contrary, if Cp were fixed, only a narrow scatter will be present. This concept can explain also the behaviour of $\overline{D}$ with respect to μ_3, μ_5, μ_9, μ_{10} and μ_{12}. In all cases, the compared measures have consistent trends: the higher the values of μ_2, μ_3, μ_5, μ_9, μ_{10} and μ_{12}, the lower the values of $\overline{D}$. In addition, $\overline{D}$ scatters with respect to all these measures, being the main cause of dispersion the value of process capability.

Then, the fraction of quality loss that can be attributed to non-ideal inspection depends on two major variables: Cp and ϕ. This is not new: process capability has been already identified as a relevant variable and ϕ has been reported as a consequence of the application of limit displacements (see §5.2).

To understand the behaviour of the remaining measures it is worth studying firstly the correlation between $\overline{\phi}$ and $\overline{\theta}$ in the simulated cases (see figure 5.9).

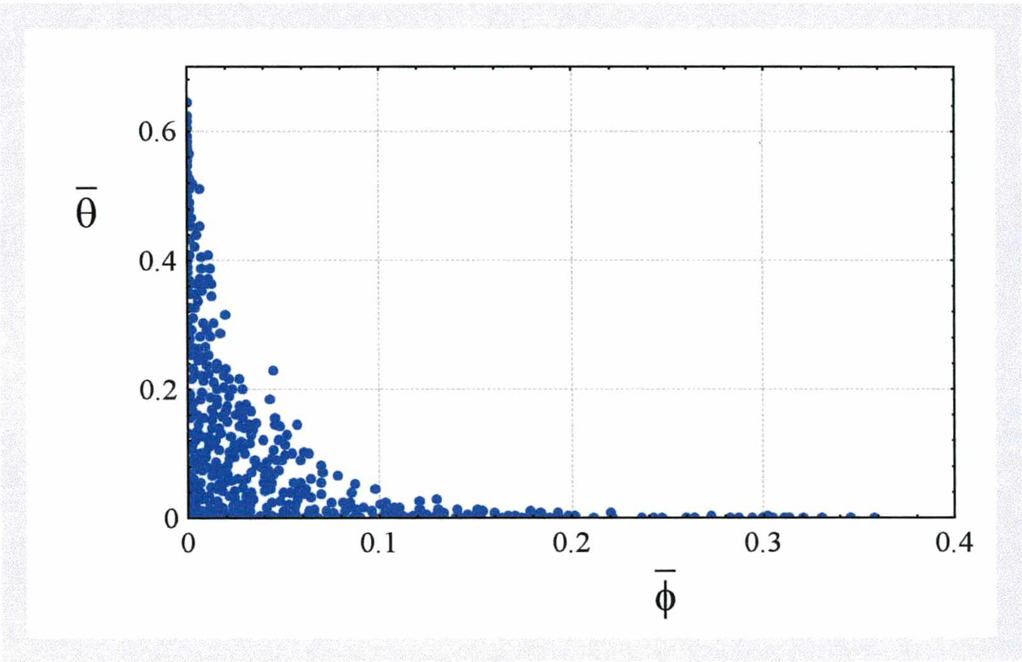

Figure 5.9: Correlation between the values of $\overline{\phi}$ and $\overline{\theta}$

(1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

Large values of ϕ are always associated with small values of θ, a situation that corresponds to inspection systems operating with large limit displacements. On the contrary, large values of θ correlate with small values of ϕ, even ϕ equal to zero. This particular behaviour is found when no limit displacements are used and measurement errors are small if compared with resolution ($\rho/\sqrt{3} \cong U_{95}$).

The possible combinations of ϕ-θ depicted in figure 5.9 permit understanding the behaviour of D against μ_1, μ_4 and μ_{11} (remember that ϕ determines the value of D).

The behaviour of D with respect to μ_6, μ_7 and μ_8 can be explained looking at the plot in figure 5.10. These measures are influenced by ϕ and θ in similar proportions. For a given value of μ_6, μ_7 or μ_8, two groups of cases can be identified. The first one is characterised by small values of θ, which can be associated with variable values of ϕ. In these cases the measure behaves as μ_1. On the other hand, small values of ϕ are associated with variable values of θ, which dominate the corresponding measure. However, in all these cases the value of the D-measure is small, forming the horizontal pattern in the plots of μ_6, μ_7 and μ_8.

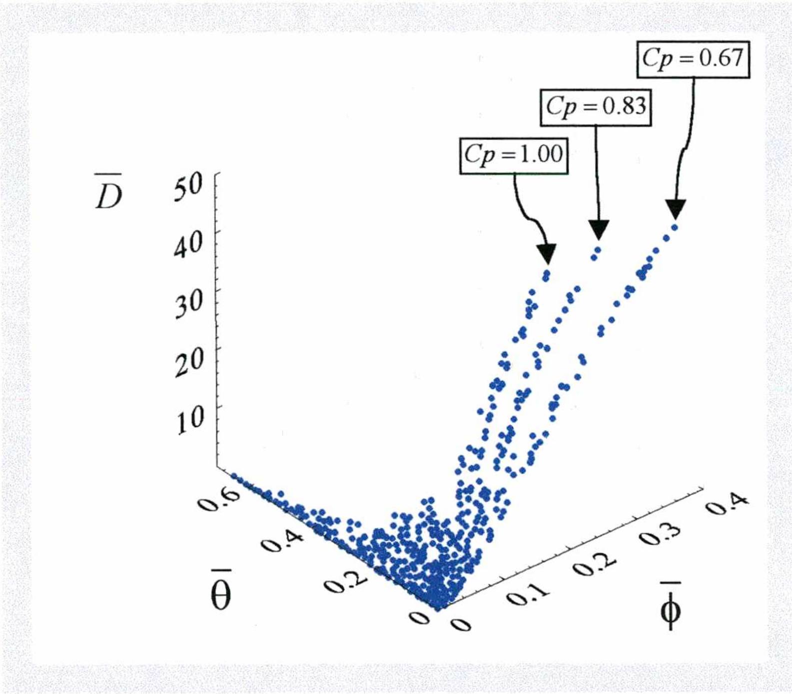

Figure 5.10: The relationship between $\bar{\phi}$, $\bar{\theta}$ and the mean value of the D-measure (1000 cases with limit displacements up to $1.5 \cdot U_{95}/T$).

At this level a question arises: does the probability of rejecting conforming units ϕ determine the value of D in the case of systems operating without limit displacements? The answer can be derived from the analysis of the plot in figure 5.11: even in the absence of limit displacements, the value of ϕ influences more the incremental quality loss due to non-ideal inspection than the value of θ.

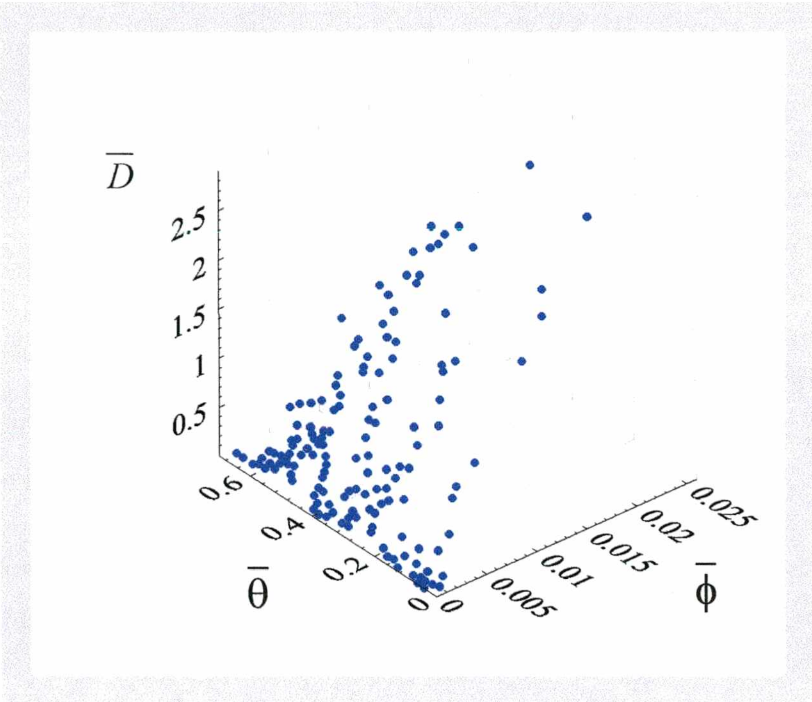

Figure 5.11: The relationship between $\overline{\phi}$, $\overline{\theta}$ and the mean value of the *D*-measure (235 cases without limit displacement).

In principle, it would be possible to obtain a regression equation expressing *D* as a non-linear function of *Cp* and ϕ. The use of this equation would allow estimating the incremental quality loss due to inspection errors in case of systems that operate with or without limit displacements. However, this solution has no practical value, because ϕ is not available in the shop floor. Estimating ϕ in real industrial problems could be almost as difficult as estimating the *D*-measure itself. No benefits can be obtained from knowing the relationship among both measures of inspection performance.

5.4 Discussion: validity of *D*-measure concept and evaluation procedure

From the analyses in §5.2 and §5.3 some general conclusions can be drawn on the validity of the *D*-measure to assess the capability of measurement systems for 100% inspection tasks. In the simulations a restricted configuration has been studied: two-sided tolerance, nominal-the-best QLF and normal manufacturing process centred on target. Nevertheless, no problems arise in applying the conclusions in this section to non-centred and non-normal processes, as well as asymmetric QLF. This is supported by the principle of independence of the limits, which states that misclassification depends on the relationship between measurement system properties and manufacturing process distribution in the neighbourhood of each specification limit. This principle is itself supported by a higher level assumption establishing that measurement errors have to be small if compared with the tolerance. Being this true, it would be impossible to misclassify a unit of a given value with respect to the two adjacent limits. As most measurement systems fulfil this condition, the analysis can be made in each limit independently. Thus, the quality loss that can be associated with non-ideal inspection is the sum of internal and external failure costs due to misclassification *in each limit*.

The same principle permits extending the conclusions of this chapter to measurement systems operating in the inspection of batches manufactured to fulfil one-sided tolerances and subjected to dimensional classification. The later case is addressed again in §6.3 by means of a numeric example.

5.4.1 Relevance of the *D*-measure as an index of inspection performance

The *fraction of the total quality loss that can be attributed to non-ideal inspection* (*D*-measure) has been proposed in §3.1 as an index of inspection performance. Due to the type of quality loss function used in its mathematical formulation, the *D*-measure is intended to assess the capability of measurement systems serving in 100% inspection or classification stations. It should not be applied to assess the capability of measurement systems used in sampling acceptance or statistical process control. In consequence, the *D*-measure is not a universal capability index, but a specific, application-oriented one. The capability statement becomes:

$$\{\text{Measurement system is capable}\} \Leftrightarrow \{D \leq D_{target}\} \tag{100}$$

where the appraiser should define the value of D_{target} to be used in the comparison. If unknown and residual systematic errors are small, the value of $\overline{D}$ can be used. On the contrary, the maximum value found in the 50 simulation loops, $\max(\hat{D}_j)$, can be used to achieve a more conservative capability assessment. The value of D_{target} will depend on the quality and cost policies of the company.

The D-measure does not belong to any of the categories of inspection performance measures listed in §1.2: neither does it evaluate the quality of measurement results nor the quality of actions on the product. Indeed, the D-measure operates directly at the level of product quality, evaluating the increment in internal and external failure costs associated with inspection errors (see figure 5.12). Consequently, it characterises an economic scale for selection and comparison of measurement systems, which is more meaningful in the industrial environment than any scale based on the quality of measurement or classification results.

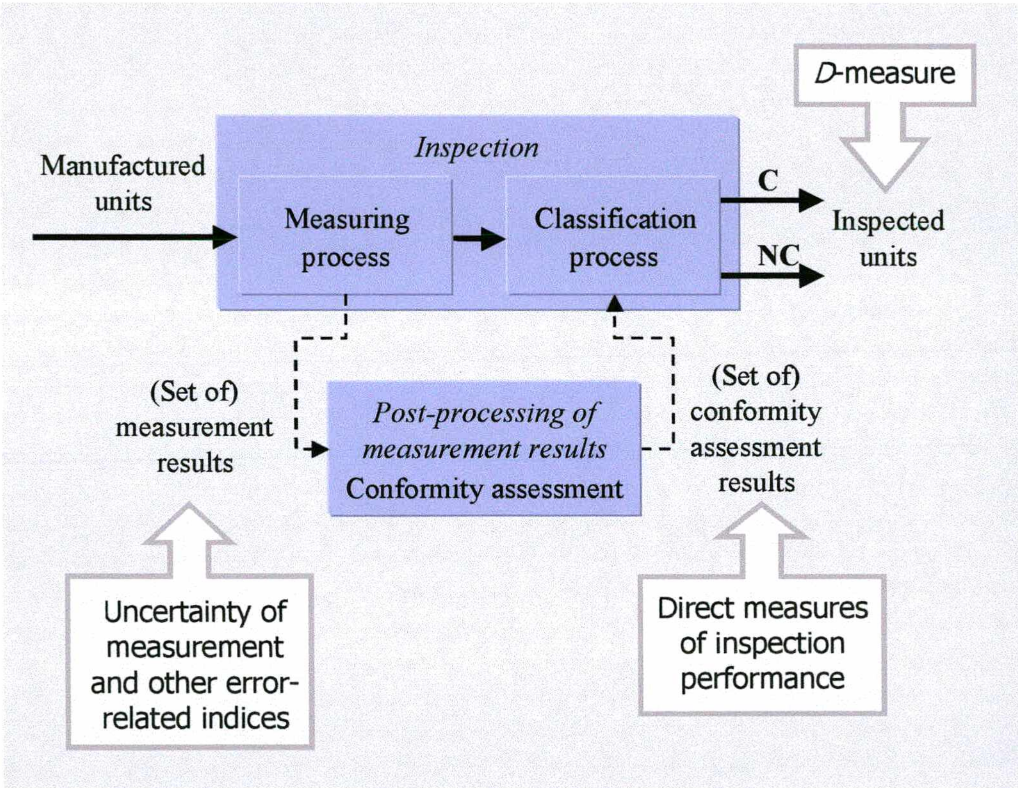

Figure 5.12: Level of action of different measures of inspection performance.

It has been shown that the scale of inspection performance defined by the D-measure is not consistent with the scales defined by other capability indices (see §5.2 and §5.3): there exists always some scatter in the cross-correlation plots. Two important factors influence this scatter. First, the D-measure does not follow the Juran's principle regarding inspection performance evaluation: it depends on the incoming quality. Indeed, the D-measure depends on the measurement system itself and metrology-related decisions, but also on design specifications, manufacturing process distribution and their interrelationships. Second, the quadratic QLF is used to weigh inspection errors instead of the step-QLF.

A remarkable fact is that the scatter of the D-measure with respect to the uncertainty per tolerance ratio is bigger than its scatter with respect to the direct measures in §2.1 (compare figures 5.2, 5.7 and 5.8). This can be understood analysing the level of action of each measure in figure 5.12. Uncertainty per tolerance ratio, and other error-related indices, refers to the quality of measurement results. On the other hand, direct measures of inspection performance evaluate the quality of actions on the product that are taken considering measurement results: they are a step farther on. Finally, the D-measure evaluates the effect of these actions on the quality of the inspected batch, being at the end of the quality control process. Each step introduces new variables, so it is natural that the biggest scatter is produced between the measures in the extremes of the chain. Between these variables, limit displacements are particularly influent. Although displacements are a consequence of measurement uncertainty, their effect appears only in the conformity assessment results. Thus, they could be responsible for the scatter between uncertainty of measurement and other downstream capability indices.

All these concepts advocate using the D-measure to assess the capability of measurement systems for 100% inspection tasks. The same reasons can be used to suggest that measurement uncertainty and other error-related indices are not meaningful for the same purpose. They should be abandoned and replaced by more specific indices, operating at the level of product quality. This statement does not preclude the use of measurement uncertainty to evaluate the capability of a measurement system for unitary measurement operations or calibration tasks.

5.4.2 Evaluation of inspection performance by computer simulation

The results earlier presented have been obtained by computer simulation, using a mathematical model of the inspection operation. This model and its implementation in the computer algorithm should be validated by experiment, to assure that simulation results are consistent with the behaviour of real inspection systems.

In principle these experiments are possible, as described in §3.3. However, the following problems could affect the reliability of the validation process:

- The experiment should be realised with real workpieces, given that it is very difficult to arrange a set of standards dense enough to sweep all the spectrum of dimensions in the manufacturing process distribution. Real workpieces present form and roughness deviations, which could result in lack of consistency among the measurement processes used to determine the *true* classification status and the *apparent* classification status.
- The number of variables of the problem is rather high. In the simplified case described in table 5.1 non-dimensional variables have been introduced. In spite of that, five independent variables remain. If a 2^n factorial experiment were designed, it would be necessary to test 32 different combinations to validate the method in the entire domain defined by the independent variables. However, by the application of the algorithm to simulated cases it has been observed that variables like limit displacement and resolution have an influence that is far from linear. This makes the 2^n design not adequate, leading to more complicated and time-consuming experimental designs.
- The algorithm computes the mean value of D, but also its uncertainty. It is known that the major contribution to the uncertainty of D is the lack of knowledge on systematic contributions to measurement uncertainty. Then, it would be necessary to perform experiments with different measurement systems presenting the same value of h but a different error curve.
- Even for 10000 units and 50 simulation loops a broad variability of D has been detected. Then, a large number of measurement operations would be necessary to reduce the uncertainty of the estimate obtained by experiment.

Due to the reasons above, no empirical validation of model and algorithm has been made in this research. Instead, the following partial validation activities have been performed /74/:

- Operational verification of the computer algorithm, to show that the transformations accurately reflect the mathematical model described in chapter 4.
- Face validation of the model in chapter 4, by discussion of the assumptions with experts in dimensional metrology.

- Face validation of the model by checking the results of its application to particular sets of inspection conditions for which the output values are known or trivial.
- Validation of the model by comparing the results with those provided by alternative methods under particular conditions. If the values of resolution and systematic errors are close to zero, the probabilities of error types I and II calculated by simulation should be equal to those computed by the method described in §2.2. It is worth remembering that this method is based in the numerical integration of the joint probability density of measurement and manufacturing errors. In consequence, it is different from the simulation procedure from the operational viewpoint.

Given the simplicity of the model, no more actions seem necessary to affirm that the proposed algorithm represents the behaviour of real inspection processes. Nevertheless, it should be remembered that coarse operator errors and behavioural patterns like flinching have been left out of the scope of this study. Despite that these errors could have an influence in the production environment, almost all theoretical analyses in metrology start from the assumption that they are negligible. Otherwise, statistical models could not be used, complicating the solution of problems like uncertainty evaluation. The results in this thesis should be used with care when human errors are predominant. On the other hand, the results can be considered representative of the behaviour of automatic inspection systems without restrictions.

6 EVALUATING INSPECTION SYSTEMS IN INDUSTRY

The aim of this chapter is to show how computer simulation can aid in the evaluation of inspection systems in industrial environment. In §6.1 the structure and capabilities of a simulation program are described. This program should not be considered as a final product, but a prototype built to test the ideas in this thesis. It makes use of the concepts in chapters 3 and 4, including also a set of routines to pre-process the input quantities. In §6.2 and §6.3, two case studies are described. The first one is about 100% inspection of bearing rings. It is based on real data and illustrates the use of the simulation procedure to analyse and improve an existing inspection task. The second case study shows how *a priori* assessment can aid in the design of a dimensional classification station. Finally, recommendations are made to support the selection and application of measurement systems in 100% inspection tasks.

6.1 Description of the prototype program

Based on the concepts presented in chapters 3 and 4, a computer simulation program has been built to evaluate the inspection performance in industrial situations. The following objectives have been established for this task:

- the program must run in a standard personal computer;
- the program should have an adequate diversity of processing options;
- input data must be available in the production environment;
- inspection performance should be expressed in terms of several alternative measures;
- the communication with the user must be visual and interactive;
- once the first simulation trial is finished, it should be possible to modify individual inputs and simulate again;
- screen output and a printed report must be provided, summarising the input parameters and the inspection performance evaluation results.

These characteristics have been embodied in the prototype program *Wininspect P.0*, built in Visual Fortran to run in Windows 95 or NT operating systems. In order to simplify the operation by a typical industrial user, the data-input interface has been separated into five thematic dialog boxes:

a) ***Initial Settings***

b) ***Inspection Condition***

c) ***Manufacture***

d) ***Measurement***

e) ***Select Measures***

The contents of these dialog boxes and the associated routines are briefly described in the following sections.

6.1.1 Processing options

The ***Initial Settings*** dialog box allows defining the inspection case to be simulated, the associated quality loss function, the probability density function for manufactured dimensions and random error and the treatment of known systematic error. The available options are listed in table 6.1.

Combination of inspection case and quality loss function	Manufacturing process distribution	Known systematic errors	Random error
• Two-sided tolerances and nominal-the best QLF • Two-sided tolerances and asymmetric QLF • One-sided tolerances and smaller-the-better QLF • Dimensional classification and nominal-the best QLF (up to 10 classes) • Dimensional classification and asymmetric QLF (up to 10 classes)	• Normal • Beta	• Corrected • Non-corrected / constant in the measuring range • Non-corrected / variable in the measuring range	• Normal • Beta • Triangular • Rectangular

Table 6.1: Processing options offered by the prototype program.

The combinations of inspection case and quality loss function are the same studied in §3.1, when deriving the equations of the D-measure. No formal restrictions have been placed on the combination of the options in the *Initial Settings* dialog box. However, the normal PDF should not be used to represent the distribution of quality characteristics for which smaller-the-better QLF is applicable. Otherwise impossible (negative) values can be generated. In that case the beta PDF should be used, adjusting the lower distribution bound to zero.

6.1.2 Routines associated with the generation of *true* values

Two dialog boxes are provided to enter the data of manufactured dimensions, one for beta-distributed processes and other for normal ones. Only one of them is presented to the user, depending on the option selected in the *Initial Settings* box. The routines associated with these boxes are displayed in figure 6.1.

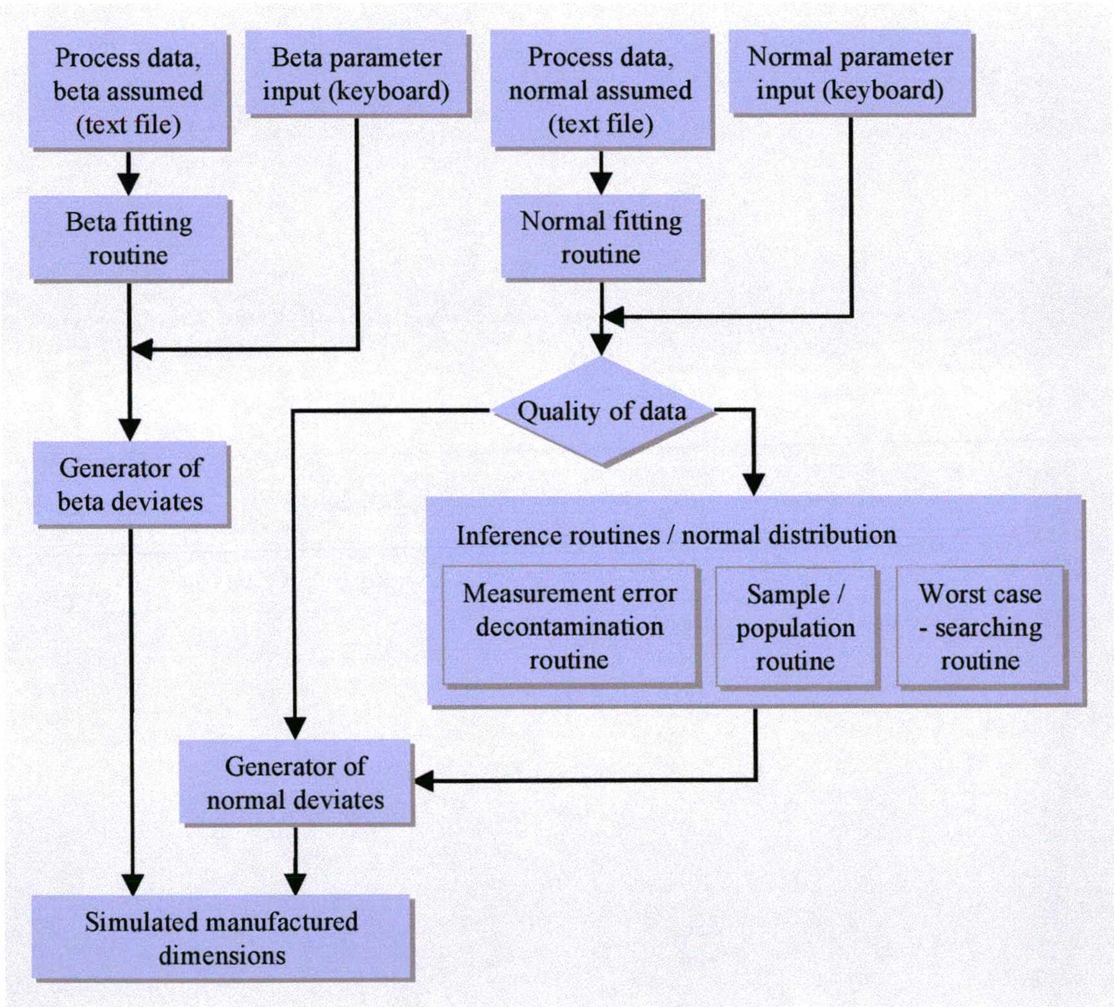

Figure 6.1: Routines associated to the generation of manufactured dimensions.

Fitting a beta PDF to data is somewhat more complicated that fitting a normal PDF, because of the number of parameters. Nevertheless, several commercial packages can fit beta probability density functions to raw data. The normal procedure is to place the lower and upper bounds of the distribution model on the minimum and maximum values of sample set. Then, the shape exponents are computed by means of simple equations /69/. Because this fitting procedure is somewhat coarse, the algorithm by J.R. He has been used in *Wininspect* /63/. It computes the lower and upper bounds of the beta model using the improved estimators by Cooke /75/, thus resulting in better fittings. The non-parametric Kolmogorov-Smirnov test has been used to evaluate the goodness of fit. Though it is suggested to have less discriminating power than the chi-square test, it can be applied directly to the data, avoiding the arbitrariness of the classification in cells. The application of these fitting and testing algorithms to simulated data has shown highly satisfactory results.

It should be remembered that the simulation routine expects parameters that describe the distribution of *true* values at a population level. Because of this, raw data used to compute the parameters should be collected with measurement systems having negligible uncertainty. In addition, the size of the sample should be big enough to minimise sampling variations (already discussed in §4.1.1).

More complete data treatment has been provided for normal processes. The input options are:

a) parameters of the population (or a large sample) of *true* values are known (*uncontaminated*);
b) parameters of a population (or a large sample) of measured values are known, being the measurement uncertainty relevant (*contaminated*);
c) *sample* parameters are known, being the measurement uncertainty negligible (*uncontaminated*);
d) *sample* parameters are known, being the measurement uncertainty relevant (*contaminated*);
e) set of raw data obtained with negligible measurement uncertainty (*uncontaminated*);
f) set of raw data obtained with relevant measurement uncertainty (*contaminated*).

Data types (b) to (f) are managed with a set of routines that compute confidence intervals for the mean and variance of the population of *true* values given the estimated parameters, the measurement uncertainty data and the sample size.

The effect of sampling and measurement errors is, in principle, not separable. To obtain the confidence intervals for the mean and variance of the distribution of *true* values given the mean and variance estimated in a sample of measured values is not straightforward. An approximate solution can be obtained analysing separately sampling effects and the contamination by measurement error.

Equations (101) and (102) describe the construction of confidence intervals for the mean and variance of a normal population, given their estimates in a sample of size *n*.

$$\mathrm{E}\{x\} = \bar{x} \pm \frac{t_{n-1;1-\alpha/2} \cdot S_x}{\sqrt{n}} \tag{101}$$

$$\frac{(n-1) \cdot S_x^2}{\mathrm{X}^2_{n-1;\alpha/2}} \leq \sigma_x^2 \leq \frac{(n-1) \cdot S_x^2}{\mathrm{X}^2_{n-1;1-\alpha/2}} \tag{102}$$

If the sample size is $n \geq 30$, the value of the *t*-Student variable that defines a two-sided, 95% confidence interval is $t_{1-n;0.975} = 1.96$. The values of the *chi*-square variable for a two-sided, 95% confidence interval can be obtained in tables. However, this procedure is not adequate for a computer program. Instead, the following equations have been used to approximate the required values of *chi*-square (see /76/):

$$\mathrm{X}^2_{n-1;0.025} = \frac{1}{2} \cdot \left[\sqrt{2 \cdot (n-1) - 1} - 1.96\right]^2 \tag{103}$$

$$\mathrm{X}^2_{n-1;0.975} = \frac{1}{2} \cdot \left[\sqrt{2 \cdot (n-1) - 1} + 1.96\right]^2 \tag{104}$$

Despite these equations are recommended when the sample size is $n \geq 100$, they also provide a reasonable approximation for $n \geq 30$ (less than 3% error with respect to the values in the table). Then, the equations of 95% confidence intervals for the mean and variance of the population are:

$$\mathrm{E}\{x\} = \bar{x} \pm \frac{1.96 \cdot S_x}{\sqrt{n}} \tag{105}$$

$$\frac{2 \cdot (n-1) \cdot S_x^2}{\left[\sqrt{2 \cdot (n-1) - 1} + 1.96\right]^2} \leq \sigma_x^2 \leq \frac{2 \cdot (n-1) \cdot S_x^2}{\left[\sqrt{2 \cdot (n-1) - 1} - 1.96\right]^2} \tag{106}$$

The confidence interval for the variance is not symmetric, unless the sample size is large ($n \geq 500$). An approximate symmetric interval can be created for the purposes of this thesis (note that the mean value of the interval is shifted with respect to the sample variance):

$$\sigma_x^2 = \frac{S_x^2}{(2 \cdot n - 6.8416)^2} \cdot \left[2 \cdot (n-1) \cdot (2 \cdot n + 0.8416) \cdot \pm 7.84 \cdot (n-1) \cdot \sqrt{2n-3}\right] \tag{107}$$

Equations (105) and (107) define a 2D confidence interval for mean and variance of a normal population given the parameters estimated in a random sample of size n. All combinations of parameters within this interval are consistent with the information content of data type (c). Type (e) data can be managed in the same way, once the estimates of mean and variance are computed for the available sample units (see equations (80) and (81)).

The distortion of the parameters of a normal population due to measurement errors has been studied by Donatelli and Schneider /19/ using the same error model proposed in this thesis. The 95% confidence intervals for the parameters of the population of *true* values are:

$$\mathrm{E}\{x\} = \mathrm{E}\{y\} \pm 0.8446 \cdot h \tag{108}$$

$$\sigma_x^2 \cong \left(\sigma_y^2 - \sigma_{ran}^2\right) \pm 0.4702 \cdot h \cdot \sqrt{\sigma_y^2 - \sigma_{ran}^2} \tag{109}$$

where h is the range within which unknown and residual systematic errors could be and σ_{ran}^2 is the variance of random errors. These equations define another 2D confidence interval. All the points within this interval represent pairs of parameters that are consistent with the information content of type (b) data.

When the available parameters have been estimated from a contaminated sample, like in data types (d) and (f), the problem becomes more complex. An approximate solution can be obtained if it is accepted that sampling variations and measurement error contamination are independent phenomena. Thus, the mean of the distribution of *true* values could be within the following confidence interval (~95%):

$$\mathrm{E}\{x\} \cong \bar{y} \pm \sqrt{\frac{(1.96 \cdot S_y)^2}{n} + (0.84 \cdot h)^2} \tag{110}$$

where $\bar{y}$ and S_y are the mean and standard deviation of the sample of measured values. The following equation describes the confidence interval for the variance of the population of *true* values:

$$\sigma_x^2 \cong \left(\frac{S_y^2 \cdot (n-1) \cdot (n+0.42)}{(n-3.42)^2} - \sigma_{ran}^2 \right) \pm \sqrt{ \left[\frac{1.96 \cdot S_y^2 \cdot (n-1) \cdot \sqrt{2 \cdot n - 3}}{(n-3.42)^2} \right]^2 + \left[0.47 \cdot h \cdot \sqrt{ \frac{S_y^2 \cdot (n-1) \cdot (n+0.42)}{(n-3.42)^2} - \sigma_{ran}^2 } \right]^2 } \quad (111)$$

In summary, three pairs of confidence intervals have been provided. The correspondence between data quality and confidence interval equations are detailed in the following table:

Data type	Mean	Variance	Data type	Mean	Variance
(a)	direct use	direct use	(d)	(110)	(111)
(b)	(108)	(109)	(e)	fit + (105)	fit + (106)
(c)	(105)	(106)	(f)	fit + (110)	fit + (111)

Table 6.2: Equations that define the confidence intervals for the parameters of the distribution of *true* values for different data qualities.

At this level a question arises: which parameter combination should be used to analyse inspection performance? In principle, any combination within the corresponding $(\mu_x; \sigma_x^2)$ domain is consistent with the available data. A conservative solution could be obtained identifying the pair of parameters that maximise the fraction misclassified by means of an optimisation routine. However, the fraction misclassified can not be known unless the simulation program itself is executed during the optimisation procedure. The time needed for this operation can not be justified in a program for industrial use. In consequence, an alternative optimisation objective has been chosen: the maximisation of the sum of probability densities in the specification limits. This procedure is illustrated in figure 6.2, for a classification operation with 5 class limits.

The parameters of a normal distribution that maximises the sum of densities in the specification limits are expected to determine the *worst inspection performance* that is consistent with the data. This occurs because the amount of probability mass concentrated in the neighbourhood of a specification limit is the most influent manufacturing process variable in the fraction misclassified.

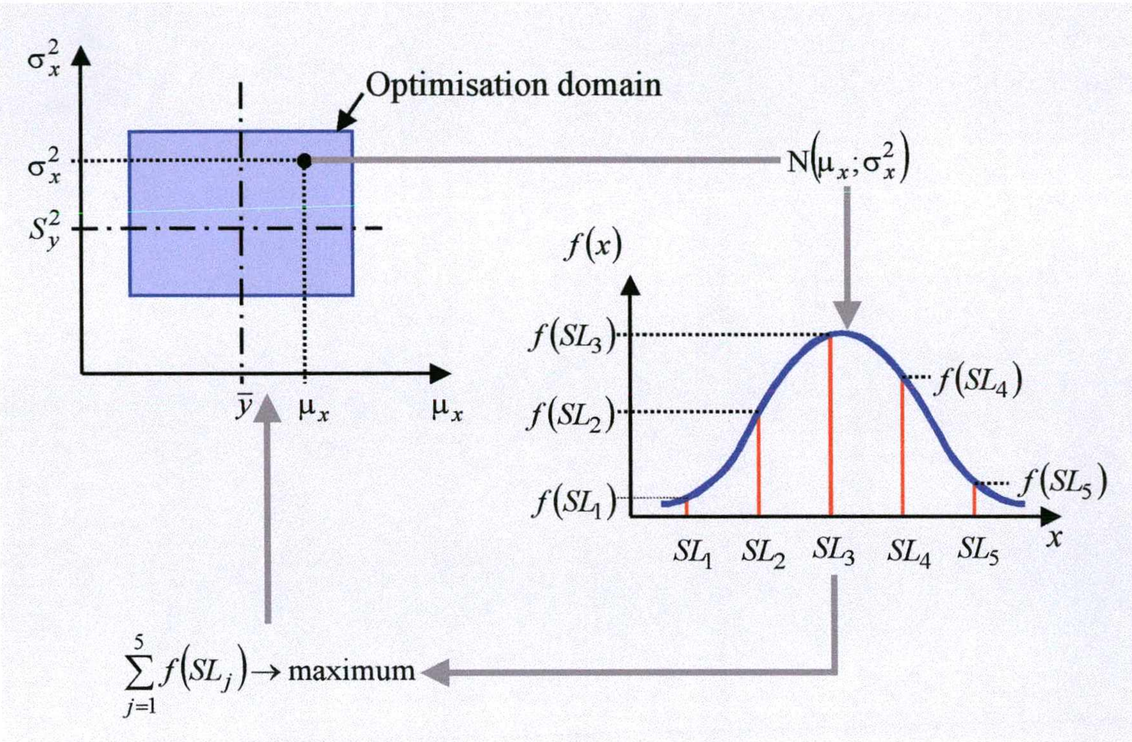

Figure 6.2: Searching for the parameters that determine the worst inspection performance.

The optimisation routine used to search for the parameters uses a quasi-Newton method and a finite-difference gradient (for more details see routine BCONF / DBCONF, in Fortran's IMSL Math Library /77/).

This searching procedure should not be used if $U_{95}/T \leq 0.2$, otherwise the increment in the fraction misclassified obtained with the optimised parameters will be hidden by sampling variations. In that case it is better to use the estimated parameters directly. It is worth commenting that, in some cases, the worst inspection performance results from displacing the mean of *true* values as much as possible within the domain of parameters. This seems not to be natural, but is the correct mathematical solution to the problem. Anyway, the option on whether optimised parameters should be adopted or not is set free to the user. Thus the empirical knowledge of the quality of estimated data can be prioritised if adequate.

6.1.3 Routines associated to the simulation of measurement results

The interrelationships among the routines associated to the generation of measurement errors and computation of measurement results are depicted in figure 6.3.

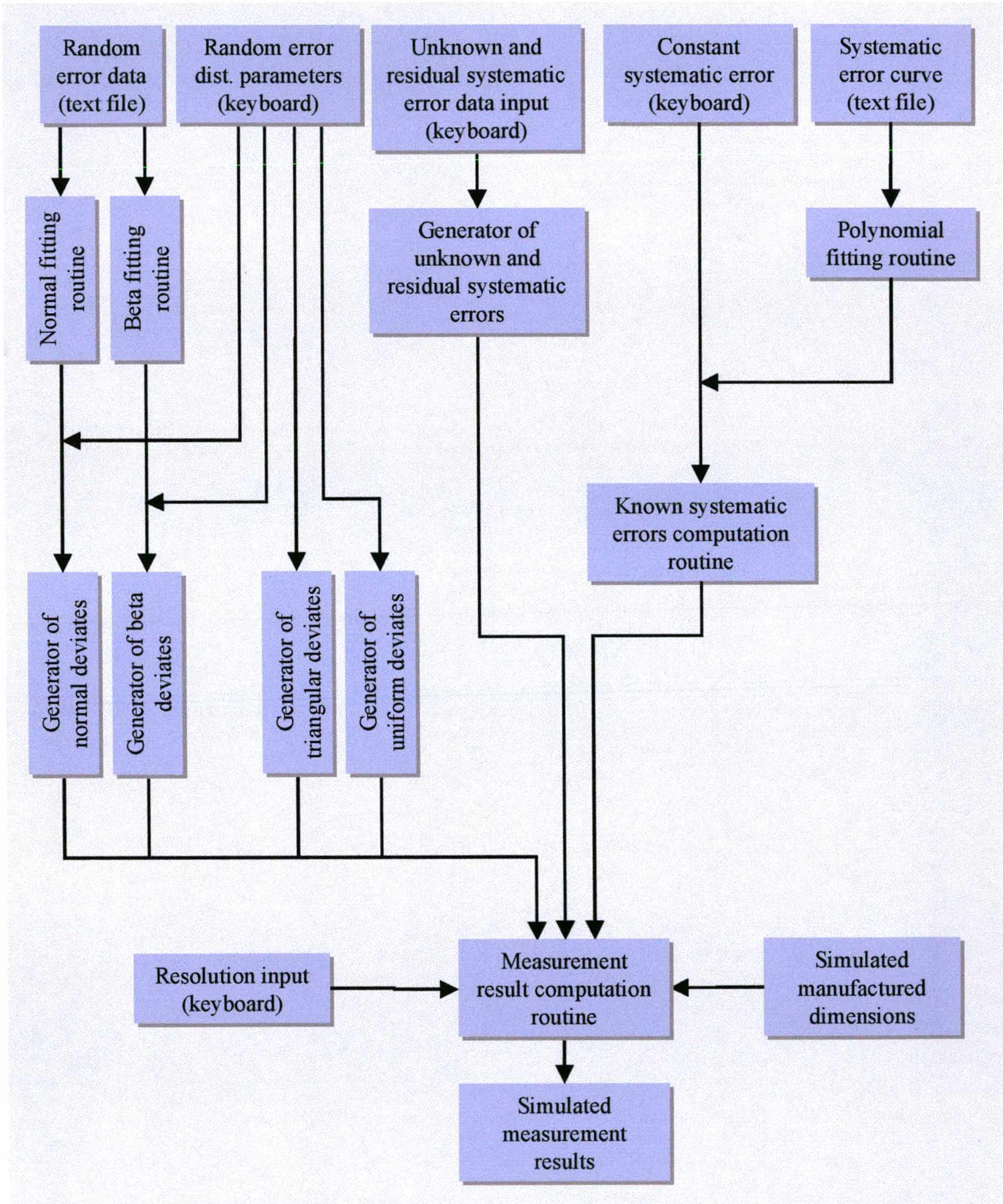

Figure 6.3: Routines associated to the generation of measurement errors and result.

These procedures implement the mathematical model of error described in §4.1 and would not need to be discussed here. However, there are some peculiarities introduced in *Winninspect* that deserve some brief comments.

It should be noted that the beta PDF has been also offered to model random measurement errors. It is not clear if this option is useful in dimensional metrology problems, because the beta variable is always positive. It has been included because the routines were already included in the program. So it is not costly to leave the option available for a hypothetical case in which the measurement error is not gaussian and always bigger than zero. The use of the normal, triangular and rectangular distribution for random errors has been already discussed in §4.1.2.

The polynomial fitting routine expects systematic error data in the shape of a two-column text file. The left column has to contain the values in the measuring range in which the calibration has been performed. The right one has to inform the corresponding systematic error values. It is important that only one systematic error value is reported for each point in the measuring range. In case several repetitions are available, the arithmetic mean should be reported. The equation of polynomial and the value of R^2 are presented to the user after the fitting procedure. This information can be used to evaluate the quality of fit.

6.1.4 Classification and inspection performance evaluation routines

The routines in this program block implement the equations in §4.2 and compute several measures of inspection performance. The equations of D-measure can be found in §3.1. Other implemented measures are described in §2.1 and §2.2. The interrelationships among these routines are depicted in figure 6.4 and do not need further explanations.

Among the measures of inspection performance described in §2.1, only those due to McCornack /37/, Youden /39/, Nelson /40/, Tiemstra /41/, Green /42/, Goodman /43/ and Cohen /44/ are computed by the program and included in the report. This selection is arbitrary and there would be no problems to compute all the measures if it is decided to do so. These measures are not available in the case of dimensional classification, because the concepts of *conforming* and *non-conforming* are not applicable.

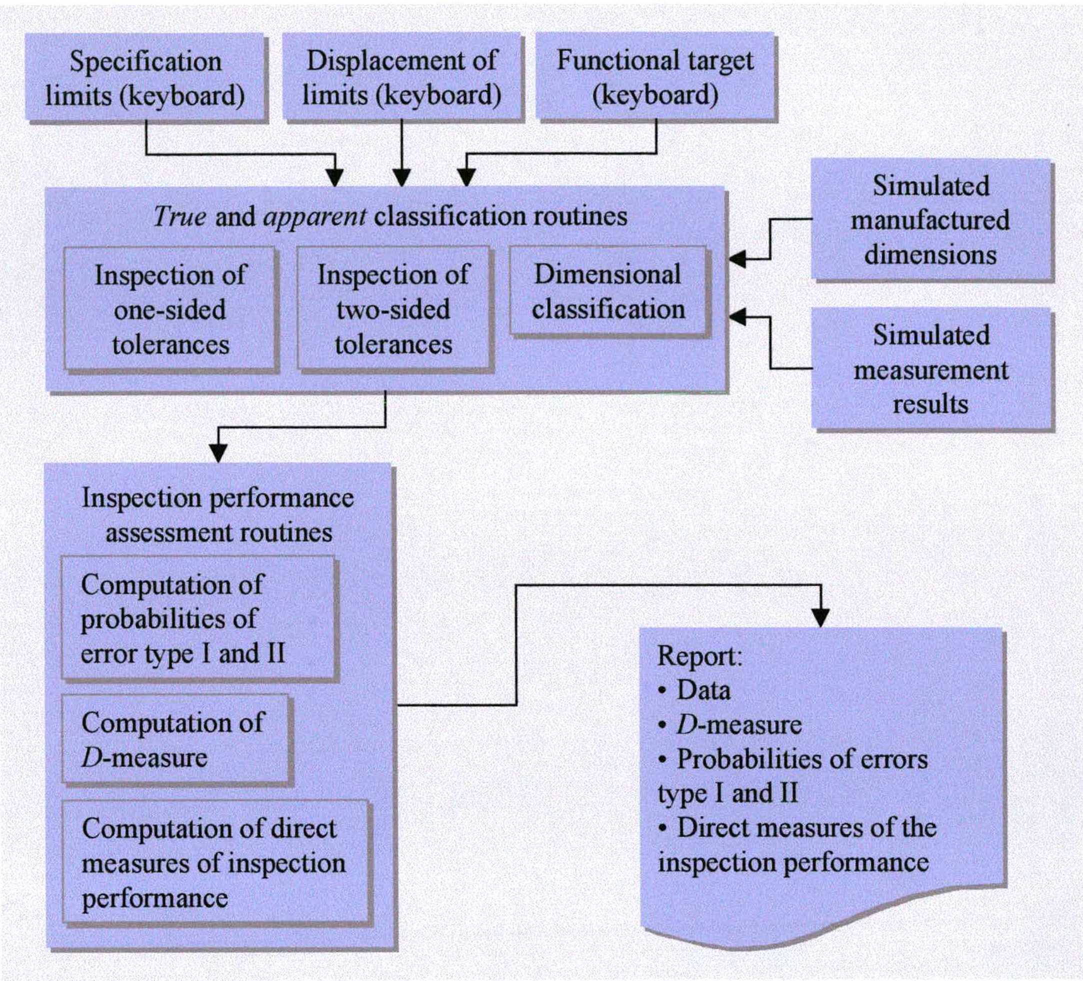

Figure 6.4: Routines for classification and inspection performance assessment.

6.1.5 Presentation of results

The mother window of the program has been used to present the simulation data and results to the user (see figure 6.5). It has been divided into three rectangular fields: the left for input data, the upper-right for a plot showing the PDF of quality characteristic and the specification and acceptance limits (to scale) and the lower-right for the inspection performance measures. This window can be printed to generate a simulation report.

In sake of simplicity and readability, only mean values of the measures of inspection performance are reported in this window. Some denomination changes have been made to simplify the understanding of a typical industrial user:

- *relative quality loss*: *D*-measure

- *total quality loss*: non-dimensional quality loss per unit including the effects of non-ideal inspection (λ^r).

The *total quality loss* is expressed in percent of A_0, cost of replacing or repairing a defective unit. The *relative quality loss* is expressed in percent of the *total quality loss*.

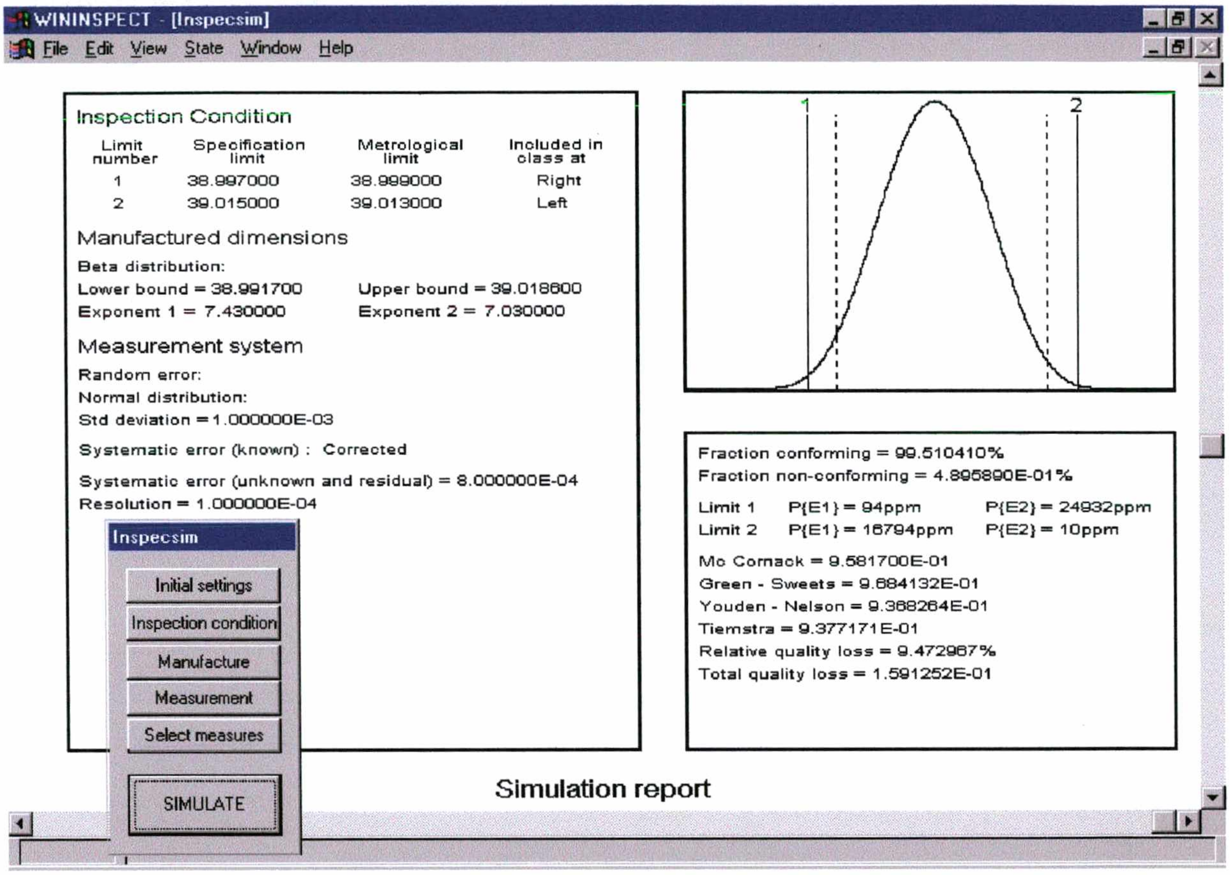

Figure 6.5: The control box and mother window of *Wininspect P.0*.

A second window has been provided with more detailed results. In it, the mean, minimum and maximum values of the *D*-measure and the probabilities of inspection errors are reported. This window can be also printed as an extended report.

6.2 Case study #1: two-sided tolerances

6.2.1 Definition of the problem

The prototype program has been applied to evaluate the inspection of the bore of hub bearing inner rings in the INA Company (Brazil). The inspection is performed over 100% of manufactured parts in automatic measuring equipment, immediately after finishing the surface by grinding. The geometrical specification of feature is rather not consistent with the current standards, but consistent with measurement method. The scheme of the measuring device is shown in figure 6.6:

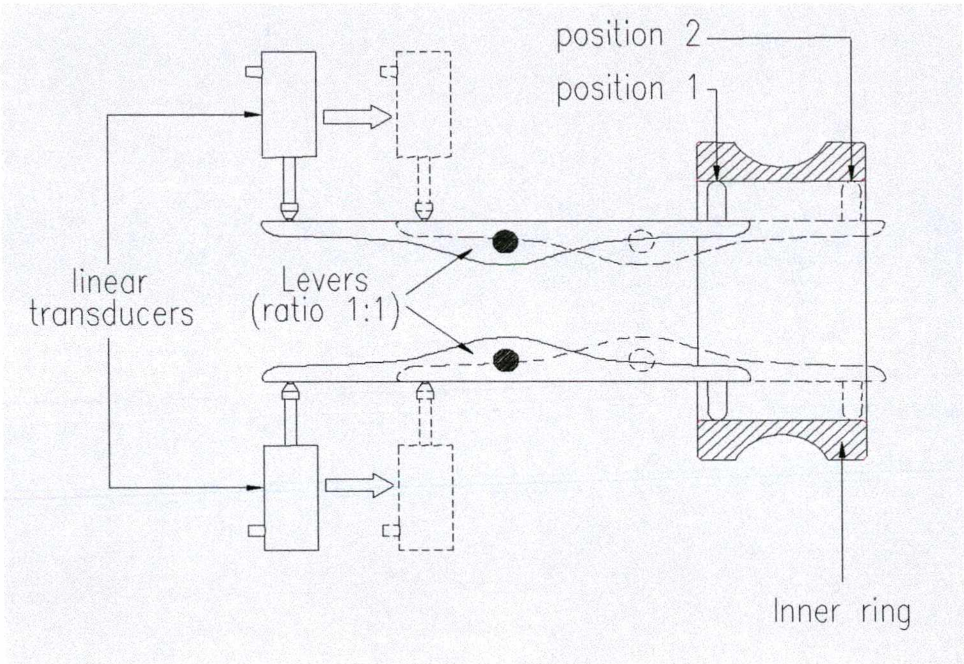

Figure 6.6: Geometry of the analysed measuring device.

The diameter is measured as a point-to point distance. During measurement, the part is rotated, sampling a first set of diameters in position 1. After that, the measuring head is displaced in axial direction to position 2, where a second set of diameters is sampled. From the collected data, three values are computed and compared with the corresponding specifications (see table 6.3):

Quality characteristic	Definition of measurand	Specification	
		Nominal	Tolerance
Diameter	Mean of point-to-point distances measured in positions 1 and 2	39mm	-3 µm / +15 µm
Circularity deviation	Maximum difference among point-to-point distances	0	0 µm / +6 µm
Conicity deviation	Difference of mean diameters computed for positions 1 and 2	0	± 5 µm

Table 6.3: Specifications and definition of measurand for the inner cylindrical surface.

Three conditions have to be fulfilled simultaneously in order a unit (bore) can be classified as conforming:

$$\left\{\begin{matrix}\text{Bore is}\\ \text{conforming}\end{matrix}\right\} \Leftrightarrow \{38.997 \leq \varnothing \leq 39.015\} \wedge \{\text{Circularity} \leq 6\mu m\} \wedge \{|\text{Conicity}| \leq 5\mu m\} \qquad (112)$$

If one of these conditions is not satisfied, the unit is qualified as non-conforming. It should be noted that the expression (112) defines three separate conformity assessment operations that could be assumed statistically independent. None of them are consistent with standard recommended practices, because:

- according to ISO 8015, all the sampled diameters -and not its mean- have to be within the tolerance interval for the workpiece to be acceptable /78/;
- circularity, as computed from sampled diameters, is not consistent with ISO 1101 /79/;
- conic deviation, as defined, does not exist in ISO 1101 standard.

Consequently, this set of specifications and measured values should be seen as a company practice, based in previous knowledge of what has to be inspected to assure the functional quality of the product. It might be observed that the comparisons determining the dimensional quality of each ring (bore) are made between consistent quantities. This makes possible the unrestricted application of simulation procedure, which is based on number-to-number comparisons.

The current study is focused on the inspection of diameter. The conformity assessment of circularity and conicity can be analysed in a similar way, but are not included in this thesis.

6.2.2 Input data

The distribution of diameters (1000 measured units) has been obtained from the SPC module of the inspection equipment, in the form of a frequency table (see table 6.4):

Deviation from nominal diameter [μm]	Relative frequency	Deviation from nominal diameter [μm]	Relative frequency	Deviation from nominal diameter [μm]	Relative frequency
$\geq$ - 8.4	1	$\geq$ 0.6	115	$\geq$ 9.6	99
$\geq$ - 6.6	2	$\geq$ 2.4	194	$\geq$ 11.4	27
$\geq$ - 4.8	4	$\geq$ 4.2	197	$\geq$ 13.2	4
$\geq$ - 3.0	10	$\geq$ 6.0	160	$\geq$ 15.0	1
$\geq$ - 1.2	43	$\geq$ 7.8	142	$\geq$ 16.8	1
Sample mean (computed from raw data):				$\bar{x} = 5.56\mu m$	
Sample standard deviation (computed from raw data):				$S_x = 3.40\mu m$	

Table 6.4: Manufacturing process data and their statistical properties.

The sample is big enough (1000 units), so that a beta PDF can be fitted to data. The estimated parameters are:

- $\hat{\alpha} = 38.9917\,mm$
- $\hat{\beta} = 39.0186\,mm$
- $\hat{\lambda}_1 = 7.43$
- $\hat{\lambda}_2 = 7.03$

The quality of this fit, as evaluated by means of the Kolmogorov-Smirnov test, is slightly better than that obtained applying the normal PDF (results not shown).

Measurement uncertainty contributions have been separated according to its statistical behaviour into random and systematic. The identified contributions are the lack of repeatability, the correction of systematic error and the differential expansion due to temperature. Repeatability data are available from the try-out of the ring, a process by which the systematic error is also estimated. The try-out is performed using a master part, which is measured 300 times. For the diameter measurement, the estimate of systematic error was

0.7μm, used to correct the system zero. This procedure is not a formal calibration, because only one point of the measuring range is analysed. This introduces an additional uncertainty in the correction beyond of that associated to the measurement of the master part itself. The total error of the correction is estimated to be within a ±0.4μm interval.

Differential expansion is due to the difference of temperature between the measurement system and the workpieces. It is assumed that both temperatures can depart in ±2°C from a common environmental temperature. The variations are considered statistically independent: one depends on the coolant and grinding conditions and the other on the measurement system itself. This results in a triangular distribution for the temperature difference in ±4°C.

The error contributions and its combination are detailed in table 6.5:

Error contribution	**Nature**	**Distribution**	**Value**
Repeatability of measuring equipment	random	normal	0.6 μm (std. deviation)
Differential expansion	random	triangular	1.8 μm (half range)
Uncertainty of correction	unknown systematic	rectangular	0.4 μm (half range)
Total random contribution = $\sqrt{0.6^2 + 0.7^2} = 0.95$ μm (std. deviation of normal PDF)			
Total unknown systematic contribution = 0.4μm			

Table 6.5: Contributions to measurement uncertainty and their classification.

The resolution at which the comparison between measured values and specification limits is made is $\rho = 0.1\mu m$.

Finally, the value of the expanded measurement uncertainty has been computed according to the ISO GUM /23/, resulting in $U_{95} \cong 2\mu m$. The corresponding uncertainty per tolerance ratio is $U_{95}/T = 2/18 = 0.11$, which could be associated to a high capability. Being measurement uncertainty small, the distortion of the distribution of *true* values due to measurement errors can be considered negligible. Then, the use of beta PDF to represent the statistical behaviour of manufactured diameters is confirmed.

The data required to tune the inspection operation have been gathered and transformed in adequate inputs to the program. In the following section the results of the simulation are presented and discussed.

6.2.3 Results

The quality of accepted units should be the main concern of those who decide upon inspection procedures and equipment. As it has been stated in previous sections, the displacement of acceptance limits with respect to specification limits can avoid the contamination of the outgoing batch with non-conforming units. From the operational viewpoint, displacing the limits seems to be an easy-to-take decision. It does not require improving the manufacturing or the inspection processes, so that no investment is necessary. Nevertheless, it is well known that limit displacements result in rejection of conforming units. Then, optimising an inspection operation would require:

- bringing the risk of accepting non-conforming units to a reasonable level (even to zero if zero-defect is required by the customer);
- minimising the loss due to rejection of conforming units.

Several simulations have been made to determine the inspection performance in presence of limit displacements. In all these cases, the displacement produced in the LSL is equal to that produced in the USL changed in sign:

$$LAL - LSL = -(UAL - USL) \tag{113}$$

The displacements have been varied from zero to 1.5 times the value of measurement uncertainty. The units which measured values are equal to the acceptance limits have been considered acceptable. The results of these simulations are depicted in figures 6.7, 6.8 and 6.9.

In figure 6.7 the values of the D-measure [%] are plotted against the absolute value of limit displacements [μm]. Square marks with continuous line represents the arithmetic mean $\overline{D}$. Upper and lower dashed lines represent the higher and lower values of $\hat{D}_j$ obtained in 50 simulation runs. All the values of D between these lines are consistent with the available knowledge.

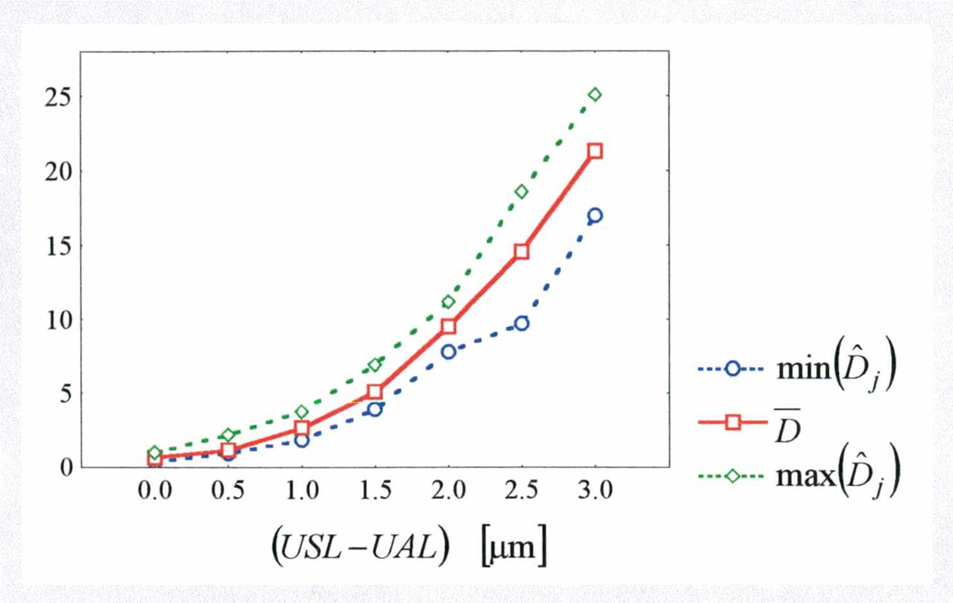

Figure 6.7: The behaviour of the D-measure (in %) for different limit displacements.

When limit displacements are zero, the percentage of the total quality loss that can be attributed to inspection errors is small, almost negligible ($0.34\% \leq D^{disp=0} \leq 0.99\%$). The application of bigger displacements makes the relative quality loss grows. For example, when displacements reach the value of measurement uncertainty the influence of the inspection system on the total quality loss is $7.75\% \leq D^{disp=2} \leq 11.16\%$, which is relevant. This increment in the value of the D-measure makes the non-dimensional quality loss with real inspection λ^r grows. This behaviour is depicted in figure 6.8. Circular marks (continuous line) correspond to the mean value of the loss with real inspection $\overline{\lambda^r}$, provided by the simulation program. Square marks (dashed line) show the behaviour of non-dimensional quality loss with ideal inspection $\overline{\lambda^i}$. The program does not provide the values of $\overline{\lambda^i}$. They have been computed from the value of $\overline{D}$ and the total non-dimensional quality loss $\overline{\lambda^r}$ applying the following equation:

$$\overline{\lambda^i} = \overline{\lambda^r} \cdot \left(1 - \overline{D}/100\right) \qquad (114)$$

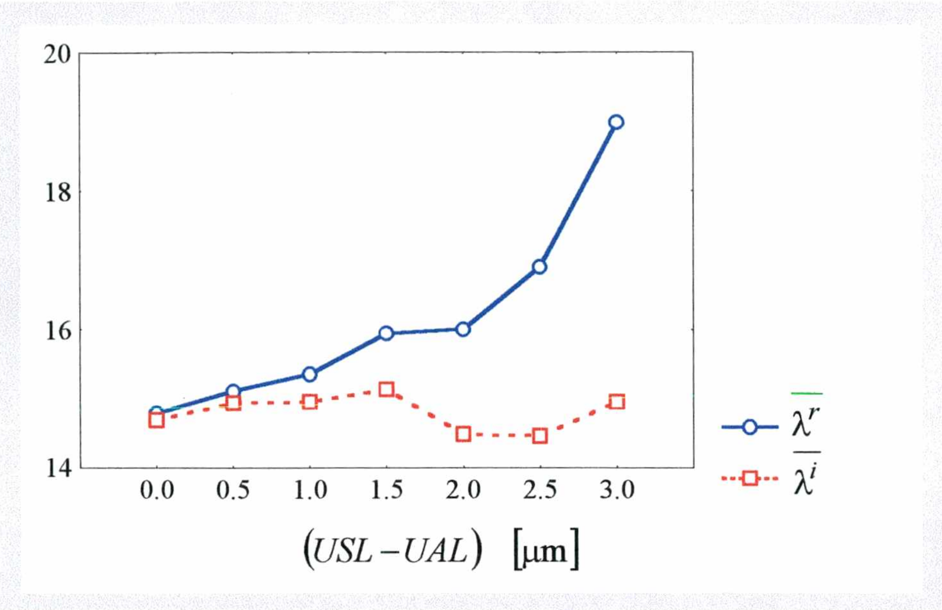

Figure 6.8: Non-dimensional quality losses with real and ideal inspection, $\overline{\lambda}^r$ and $\overline{\lambda}^i$ respectively (in %), for different values of limit displacements (in μm).

It can be observed that $\overline{\lambda}^i$ remains approximately constant when the effect of inspection errors is subtracted. Thus the difference $\left(\overline{\lambda}^r - \overline{\lambda}^i\right)$ is the contribution of measurement system to the quality loss.

A more comprehensive analysis can be performed if the probabilities of inspection errors corresponding to the same simulation runs are plotted against limit displacements. Figure 6.9 shows the behaviour of probabilities of error type I and II in both specification limits. In the first row the probabilities of error type I (in parts per million - ppm) are plotted against limit displacements (in μm). The probabilities of error type II are plotted in the second row. The two graphs on left correspond to LSL and the two on right to USL. Just like the *D*-measure, the probabilities of inspection errors have been reported as dispersion ranges. These ranges include all the values obtained in 50 simulation loops, according to equations (87) and (92).

The probabilities of inspection errors are both higher in the LSL than in the USL, regardless the value of limit displacements. This can be easily justified looking at the plot in the simulation report (see figure 6.10). The slight skewness of beta PDF and the deviation of the mean with respect to the middle of tolerance interval caused the probability mass in the neighbourhood of LSL to be greater than in the neighbourhood of USL. As the proportion misclassified increases with the probability mass involved, the probabilities of inspection errors will be higher in LSL than in the USL. This behaviour is consistent with simulation results in figure 6.9.

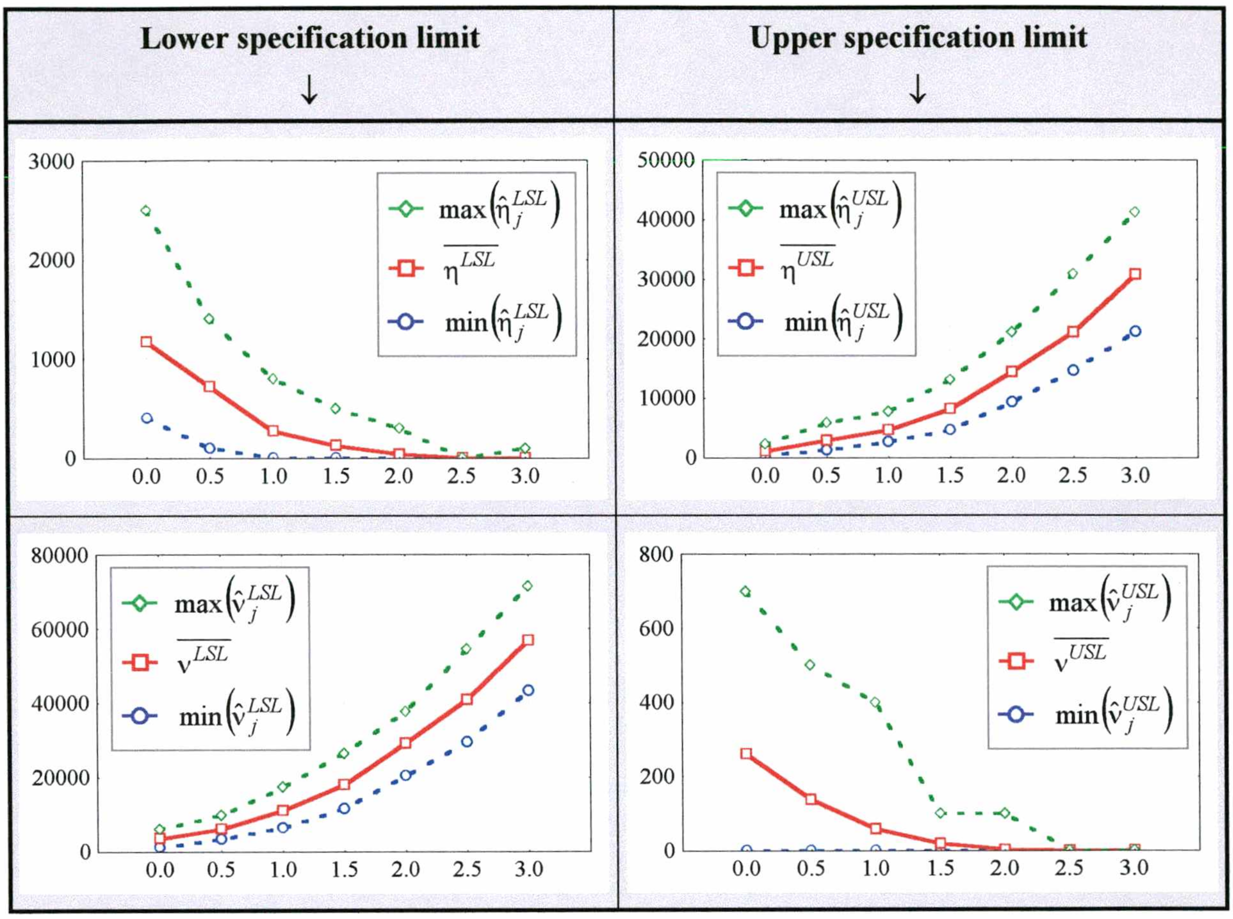

Figure 6.9: Probabilities of errors type I and type II (in ppm) for different limit displacements (in μm).

It can be observed that the probability of rejecting conforming units is higher than the probability of accepting non-conforming ones in both specification limits, even if limit displacements are zero. This is also consistent with the theory because the resolution of measurement system is small if compared with the measurement uncertainty ($\rho/U_{95} = 0.1\mu m/2\mu m = 1/20$). If the resolution were coarser, it would increase the probability of accepting non-conforming units, reducing the probability of rejecting conforming ones at the same time.

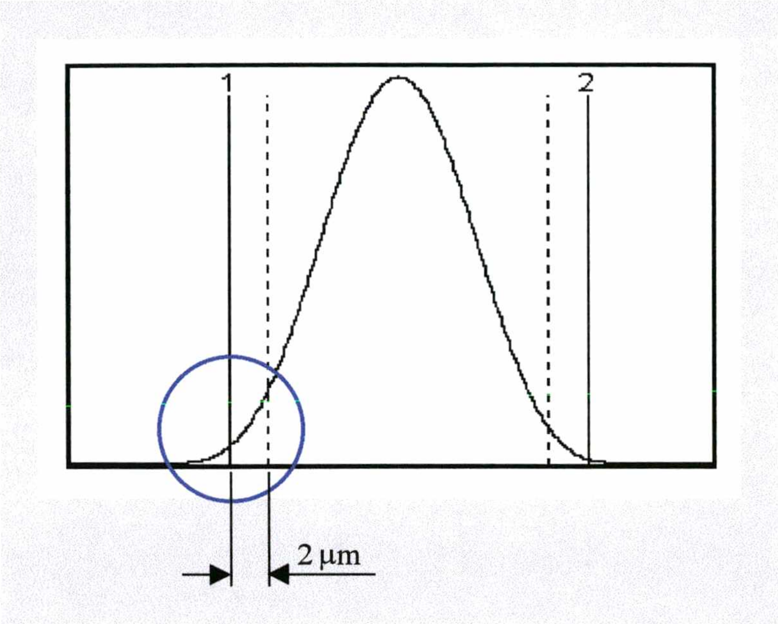

Figure 6.10: Beta PDF, specification limits and acceptance limits (displacements: 2μm)

Finally, it can be observed that the displacement of the limits should be at least 2.5μm (or $1.25 \cdot U_{95}$) to achieve zero-defects. The consequence of the application of such displacements is an increment in D from $0.37\% \leq D^{disp=0} \leq 0.99\%$ to $9.65\% \leq D^{disp=2.5} \leq 18.55\%$, with the corresponding increment in the total quality loss.

In summary, the following concepts should be considered to tune the inspection operation:

- the lowest acceptable value of D is obtained when the acceptance limits are placed on the specification limits (zero displacements);
- major contributions to D come from misclassification in the neighbourhood of LSL.

Inspection tuning has to consider the function of the part to be inspected. Before deciding the value of limit displacements it is worth evaluating if the benefits of zero-defect approach can justify such increment in the total quality loss. It should be remembered that non-conforming units accepted by mistake in the LSL and USL could not be equally harmful to the quality of outgoing product. In the analysed case the acceptance of quality characteristics which values are $x < LSL$ must be avoided. On the contrary, the acceptance of a small number of units with $USL < x < (USL + 1\mu\text{m})$ is not so prejudicial to quality. This is similar to an extension of tolerance interval. However, not all units presenting diameters in the extension range will be accepted: only those that are misclassified by the instrument.

Based on the constraints above a first proposal can be made regarding limit displacements. There is not so much to do in LSL. The zero-defect requirement forces the adoption of a 2.5μm displacement, so that the acceptance limit becomes $LAL = 38.9995\mu m$. In USL the displacement can be reduced to -1.5μm. This permits fulfilling the requirement on the extended tolerance and maximising the D-measure at the same time. The upper acceptance limit becomes $UAL = 39.0135\mu m$. The simulation results can be observed in table 6.6.

	Mean value	**Dispersion interval**
D-measure	10.90%	$8.37\% \leq D \leq 13.12\%$
Accepting non-conforming units in LSL	4ppm	$0ppm \leq \eta^{LSL} \leq 99ppm$
Rejecting conforming units in LSL	40222ppm	$30099ppm \leq \nu^{LSL} \leq 49699ppm$
Rejecting conforming units in USL	8078ppm	$4800ppm \leq \eta^{USL} \leq 11800ppm$
Accepting non-conforming units in USL	10ppm	$0ppm \leq \nu^{USL} \leq 99ppm$

Table 6.6: Results of inspection performance evaluation for the first proposal.

The requirement of zero-defect can be considered fulfilled in the LSL, though the probability of error type I is not identically equal to zero. The small found values could be attributed to the tails of the normal PDF used to represent random measurement errors, so that they can be ignored. In the USL, the probability of accepting non-conforming units is similar to that found in LSL, though the limit displacement was only -1.5μm. This is because the lower concentration of probability mass in the neighbourhood of the USL.

The use of limit displacements to preserve the quality of outgoing batch is not cheap. On average, 4.8% of the total inspected will be rejected by mistake. This makes the value of the D-measure to be relatively high: up to 13% of the total quality loss can be attributed to inspection system. In the opinion of the author this solution is not satisfactory. However, nothing better can be done if the quality requirements above have to be fulfilled.

If the zero-defect requirement is removed in the LSL a more economical manufacture can be achieved. Reducing the displacement in this limit to 1.5μm ($LAL = 38.9985\mu m$), the probability of rejecting conforming units will decrease drastically, at the expense of a slight increase in the probability of accepting non-conforming ones. This small number of non-conforming units will present dimensional deviations with respect to the LSL that are smaller than 1.0μm ($38.996mm \leq diameter \leq 38.997mm$). A similar concept can be used in the USL, reducing the displacement to -1.0μm ($UAL = 39.0140\mu m$). Units accepted by mistake in this limit will have diameters in the interval $39.015mm \leq diameter \leq 39.0165mm$, which represents an exiguous tolerance enlargement. The application of the re-defined displacements leads to the simulation results in table 6.7.

	Mean value	**Dispersion interval**
D-measure	4.16%	$2.97\% \leq D \leq 6.01$
Accepting non-conforming units in LSL	71ppm	$0ppm \leq \eta^{LSL} \leq 399ppm$
Rejecting conforming units in LSL	18037ppm	$11900ppm \leq \nu^{LSL} \leq 26100ppm$
Rejecting conforming units in USL	4434ppm	$2000ppm \leq \eta^{USL} \leq 7100ppm$
Accepting non-conforming units in USL	81ppm	$0ppm \leq \nu^{USL} \leq 399ppm$

Table 6.7: Results of inspection performance evaluation for the second proposal.

Note the drastic reduction in the value of the D-measure with respect to that in table 6.6. On the contrary the total number of parts accepted by mistake have increased to 0.08% of the total inspected batch, that is not so much indeed.

Using equations (24) and (25) the values of ϕ and θ can be estimated for the probabilities of inspection errors in table 6.7. As informed in the simulation report, $q = 0.5\%$ and $p = 99.5\%$. Then, the expressions for the worst-case intervals are $1.4\% \leq \phi \leq 3.3\%$ and $0.0\% \leq \theta \leq 16.0\%$. The mean values of ϕ and θ can be calculated in a similar way: $\overline{\phi} = 2.2\%$ and $\overline{\theta} = 3.0\%$. From these values, the indices of inspection performance in table 2.1 can be computed. Several of them are already informed in the simulation report of *Wininspect P.0*.

These calculations close the analysis of Case #1. Two more studies on the performance of the same inspection system can be found in /80/: effect of an uncorrected systematic measurement error and effect of instability in the manufacturing process distribution mean. These studies are prior to the development of the *D*-measure, so that they use the probabilities of errors type I and type II as measures of inspection performance.

6.3 Case study #2: dimensional classification

This is an artificial case study on dimensional classification. It has been included to show the characteristics of *a priori* evaluation of classification performance. Data are feasible, but do not correspond to any particular (real) manufacturing operation or measurement system.

A classification facility has to separate a batch of manufactured units in eight dimensional classes of 10μm width (see data in table 6.8).

Data group	Data description	Values
Classification condition	Class limits	$SL_1 = 86.010mm$ $SL_2 = 86.020mm$ $SL_3 = 86.030mm$ $SL_4 = 86.040mm$ $SL_5 = 86.050mm$ $SL_6 = 86.060mm$ $SL_7 = 86.070mm$
	Limit displacements	0.000mm (all)
	Limits included in class at	right (all)
Manufacturing process	Normal PDF parameters	$\mu_x = 86.040mm$ $\sigma_x = 0.01mm$
Measurement system	Std. deviation of random errors	$\sigma_{ran} = 1.0\mu m$
	Unknown and residual systematic errors	$h = 1.0\mu m$
	Resolution	$\rho = 0.1\mu m$
	Uncertainty of measurement	$U_{95} = 2.2\mu m$

Table 6.8: Data to simulate the classification operation.

Units in the extreme classes are considered non-conforming: they must not be delivered to assembly. No limit displacements is to be applied to correct the balance between type I and type II errors. All the units whose measured values are equal to specification limits are classified in the class to right of the corresponding limit. Based on previous knowledge, the manufactured population (*true* values) is supposed to be normal, with standard deviation equal to the class width.

An automatic measuring machine is designed to perform the classification task. Estimated metrological characteristics of the system are also informed in table 6.8. Note that the uncertainty per tolerance ratio is $U_{95}/T = 2.2/10 = 0.22$, so the system could be considered scarcely capable. The relationship between U_{95} and resolution is $U_{95}/\rho = 2.2/0.1 = 22$. Then, rounding up will not affect the classification behaviour relevantly /59/.

The values of the D-measure obtained in the 50 simulation runs (10000 units) are within the interval $10.28\% \leq D \leq 17.35\%$, being the arithmetic mean $\overline{D} = 13.22\%$. The probabilities of error type I and type II are reported in table 6.9.

The table makes evident that inspection performance depends on the position of the class limit within the scatter range of manufacturing process. Class limits in the middle of the range are subjected to heavy misclassification. It should be remembered that the probabilities of type I and type II errors are referenced to the total number of classified units. If the values of ϕ and θ were computed for each class independently, the variations of these indices among classes would be smaller. This is because the classes that are subjected to higher values of η and ν have also more probability mass (see equations (24) and (25)).

The quality loss associated to each manufactured unit is, *on average*, 39% of A_0 (cost of replacing or repairing a non-conforming unit). This high value is due to the presence of class limits in regions with high probability density of manufactured units. The same reason is valid to justify the value of the D-measure: the incremental quality loss due to non-ideal inspection is, *on average*, 5% of A_0 per unit.

This performance can be acceptable or not, depending on the design requirements. In case it were necessary to improve the performance, the main resource would be to improve measurement system properties. For example, reducing random and residual systematic errors to $\sigma_{ran} = 0.5\mu m$ and $h = 0.5\mu m$, the values of the D-measure obtained in the 50 simulation runs (10000 units) are within the interval $2.70\% \leq D \leq 5.57\%$, being the arithmetic mean $\overline{D} = 3.56\%$. The total quality loss per unit dropped to 35 % of A_0. This value is still high, but it should be remembered that, even in the absence of classification errors, the total quality loss in batches subjected to dimensional classification is high. Indeed, the total quality loss per unit with ideal classification is 33.7% of A_0.

	Error type I	**Error type II**
$SL_1 = 86.010\text{mm}$	$\overline{\eta}^{SL_1} = 154\text{ppm}$	$\overline{\nu}^{SL_1} = 246\text{ppm}$
$SL_2 = 86.020\text{mm}$	$\overline{\eta}^{SL_2} = 2238\text{ppm}$	$\overline{\nu}^{SL_2} = 2884\text{ppm}$
$SL_3 = 86.030\text{mm}$	$\overline{\eta}^{SL_3} = 10840\text{ppm}$	$\overline{\nu}^{SL_3} = 11132\text{ppm}$
$SL_4 = 86.040\text{mm}$	$\overline{\eta}^{SL_4} = 18934\text{ppm}$	$\overline{\nu}^{SL_4} = 17438\text{ppm}$
$SL_5 = 86.050\text{mm}$	$\overline{\eta}^{SL_5} = 12322\text{ppm}$	$\overline{\nu}^{SL_5} = 9936\text{ppm}$
$SL_6 = 86.060\text{mm}$	$\overline{\eta}^{SL_6} = 2994\text{ppm}$	$\overline{\nu}^{SL_6} = 1934\text{ppm}$
$SL_7 = 86.070\text{mm}$	$\overline{\eta}^{SL_7} = 256\text{ppm}$	$\overline{\nu}^{SL_7} = 136\text{ppm}$

Table 6.9: Results of classification performance evaluation.

The example above is enough to illustrate how the computation of the D-measure can aid in the improvement of dimensional classification operations. It has also been an example of the use of the simulation program based on *a priori* data.

6.4 Discussion: the use of measurement systems in 100% inspection

In the preceding case studies, the D-measure has been applied together with the probabilities of error type I and type II to improve the global performance of inspection operations. Although most of the concepts developed in §5.2 and §5.3 have been clarified by their application to the cases in this chapter, it is worth reviewing them here in a more general form. This will support basic recommendations for the selection and application of measurement systems in industrial metrology.

In today's customer-supplier relationship, the contamination of a product batch with non-conforming units is considered unacceptable from the viewpoint of quality. It has been already mentioned that the most cost-effective means to reach zero-defect is to produce the quality characteristic with a manufacturing process having enough capability (i.e. $\hat{C}p \geq 1.33$ and/or $\hat{C}pk \geq 1.67$). Unfortunately, processes with insufficient capability can be found everywhere in industry. In such cases, 100% inspection must be used to identify and separate non-conforming units.

An inspection system should be capable of identifying *all* non-conforming units. Additionally, it should not classify conforming units as non-conforming, to prevent unjustified product losses. Real inspection systems can not fulfil simultaneously both requirements, due to the effect of measurement errors and resolution. Then, if the zero-defect requirement must be fulfilled, it is necessary to create a reduced acceptance interval. In particular, ISO 14253-1 (E) suggests that the supplier should reduce the tolerance interval by two times the value of measurement uncertainty to prove the conformity with specifications /28/.

However, results in §5.2 and §6.2 show that the value of the *D*-measure grows with the application of limit displacements. The preservation of the quality of the outgoing batch causes internal failure costs to increase drastically, because of massive rejection of conforming units. This affects the total quality loss of the inspected batch, deteriorating the economy of production. It should be noted that there is a fraction of the incremental quality loss directly associated with limit displacements. This loss would exist even in absence of measurement errors and, depending on the value of displacement, it can offset the desirable effect of an increase in the manufacturing process capability.

The two preceding paragraphs can be summarised in a simple statement: zero-defect by 100% inspection is feasible, but expensive. The decision of reducing the acceptance interval is easy to take, because it does not require investment or special knowledge. However, it should be treated as a compromise between appraisal and failure costs. The better the measurement system properties, the smaller the displacements needed to achieve zero-defect. Then, investment in better equipment and inspection conditions pays back in the form of fewer conforming units rejected by mistake. It is important to realise that limit displacements should be neither undersized, nor oversized. To achieve a correct assignment of displacements, it is necessary to have adequate knowledge on measurement system and manufacturing process properties. Relevant properties are (already detailed in chapter 5): probability density of manufactured units and its relationship with specification limits, measuring device resolution and measurement errors.

The results in §5.2 and §6.2 show that, depending on the probability mass concentrated in the neighbourhood of the specification limit, displacements of one measurement uncertainty could not be sufficient to reach zero-defect. The higher the probability mass, the higher the values of limit displacement that should be applied to achieve zero probability of accepting non-conforming units (up to $1.5 \cdot U_{95}$ in some cases). It is worth remembering that, in case of two-sided tolerances, the probability mass in the neighbourhood of the limits depends on the process capability. Processes with low value of $\hat{Cp}$ could require large displacements in both specification limits to preserve the quality of the outgoing batch. If the process has a skewed distribution or its mean deviates from target ($\hat{Cp} \neq \hat{Cpk}$), it could be necessary to apply different displacements in each specification limit. This criterion remains valid in case of batches manufactured to fulfil one-sided tolerances, but *CPU* and *CPL* capability indices must be used instead of *Cp*.

Measuring device resolution could also affect the value of limit displacements that are necessary to approach zero probability of classifying a non-conforming unit as conforming ($\theta = 0$). In the inspection of batches subjected to one- or two-sided tolerances, it is common practice to classify the units whose measured values are equal to the acceptance limits as conforming. Under this condition, the coarser the resolution, the higher the probability of accepting non-conforming units, so requiring larger limit displacements to fulfil the zero-defect requirement. Consequently, the resolution should be maintained as fine as possible. Indeed, the recommendation that the resolution should be smaller or equal than one-tenth of tolerance /81/ is rather insufficient. In the opinion of this author, the resolution of measuring device to serve for 100% inspection should fulfil two simultaneous conditions: $\rho \leq T/100$ and $\rho \leq U_{95}/20$. If both requirements are satisfied, the effect of rounding can be ignored.

The effects of random and systematic measurement errors on the inspection performance have been described in §5.2. No further discussion is needed regarding random errors. However, the presence of systematic errors affects drastically the inspection performance and deserves some comments. It is a metrology dogma that *known systematic errors must be corrected.* In spite of that, most systems in industry are calibrated (zeroed) in only one point of the measuring range. Afterwards, residual errors and other non-corrected long-term variations are estimated and considered as contributions to uncertainty of measurement. It should be remembered that systematic contributions to uncertainty not only affect the mean value of inspection performance measures, but also their uncertainty. To approach zero-defect it could be wise using the worst-case values within the corresponding dispersion intervals. For example, in case of two-sided tolerances the displacements should be assigned to produce $\max\left(\hat{\eta}^{LSL}\right) \cong 0$ and $\max\left(\hat{\nu}^{USL}\right) \cong 0$, which implies $\max\left(\hat{\theta}\right) \cong 0$. Thus, the diminution of systematic contributions to uncertainty will produce a double benefit, leading to smaller limit displacements. In this context, the calibration of the measurement system in its application environment becomes critical. Given that inspection errors are produced in a region around each acceptance limit, it could be enough to calibrate the system with master parts whose values are equal those limits. For example, if the specification limits were $LSL = 38.997$ and $USL = 39.015$ and the displacements were $+0.002$ and -0.002, the master parts should have $x_1 = 38.999$ and $x_2 = 39.013$. This procedure will reduce systematic errors in the region of the measuring range where they affect more the classification performance, being cheaper and simpler than a calibration in the complete measuring range. In order to minimise the effect of medium- and long-term variations, the calibration should be performed as frequently as economically feasible.

The presentation above leads to a set of recommendations about how to seek for zero-defect using 100% inspection:

- Automatic inspection systems should be preferred. If they are not available, human errors and behavioural patterns should be prevented providing the inspector with unambiguous information about how to classify each unit: conforming or not-conforming, green or red, good or no-good... Scale comparisons should be avoided.
- Inspection systems with good discrimination should be preferred (resolution $\rho \leq T/100$).
- All known systematic errors should be corrected;

- Systematic contributions to uncertainty should be minimised implementing proper calibration procedures. Systematic errors can vary in time, so measuring instruments should be calibrated periodically.
- Measurement uncertainty should be computed according to recognised procedures. Uncertainty should not be underestimated nor overestimated: the best possible use of the available knowledge should be made. A reasonable objective is $U_{95} \leq T/10$ (higher values are not encouraged).
- Considering the values of measurement uncertainty, resolution and the relative position of the process distribution with respect to specification limits, limit displacements should be estimated to satisfy the zero-defect requirement.

The recommendations above could serve as a guide for the selection and application of measurement systems in 100% inspection tasks. Once the parameters of measurement system and manufacturing process are defined, the simulation program proposed in this chapter could be used to predict the behaviour of the system. In this context, the role of the D-measure would be to show the influence of measurement system on the total quality loss. Probabilities of errors type I and type II would reveal the degree of contamination of the accepted batch with non-conforming units, so guiding the inspection refmement process. Finally, the non-dimensional quality loss λ^r would provide information on the internal and external failure costs of the complete production system, *i.e.*, manufacturing plus inspection. In this manner, the three measures contribute to a unique goal: to achieve the best possible product quality with the minimum overall cost.

After this presentation on how to achieve zero-defect by 100% inspection a question arise: is the quality improvement achieved in the outgoing batch important enough to justify the drastic increment in internal failure costs? The today opinion of the international standard's community, expressed by ISO/FDIS 14253-1:1997(E), seems to be *YES* /28/. The standard establishes rules for proving conformance and non-conformance with specifications, which apply as *default* rules in the supplier chain. In principle, it gives the same rights to supplier and customer: supplier shall prove the conformance with specification and customer shall prove the non-conformance. However, in the current market conditions, the standard is being used by the customers to force a quality improvement of supplied products at no extra cost.

As a matter of fact, ISO 14253-1 emphasises the viewpoint of the customer, who does not want to receive non-conforming units at all (see note in §6.1 of the standard).

In the opinion of this author, based on the results in §5.2 and §6.2, the requirement of zero-defect leads to unjustified supplier costs when 100% inspection is involved. In §6.2 it has been shown that the application of displacements that are smaller than those required to achieve zero-defect can lead to more economic production conditions. By this means, the probability of accepting non-conforming units can be maintained low. The *true* values of those non-conforming units accepted by mistake would be close to the specification limits, being the maximum deviations from limits predictable. This proposal seems valid. Nevertheless, it requires accepting that the specification interval can be enlarged without loss of functional performance, which is a rather conflictive decision. The following reasons advocate accepting the enlargement of tolerance interval:

- The tolerance assignment process is responsible for transforming the allowable functional variation into allowable dimensional/geometrical variation. In most cases this process is rather informal, leading to tolerances that have an unknown associated uncertainty. Knowing this fact, designers specify tolerances that are tighter than necessary, to cover themselves against product malfunction.
- In spite of the avalanche of methods and software packages to help in the statistic allocation of tolerances, company know-how is still the best way to achieve an effective assignment of tolerances. In this case know-how can be defined as the accumulation of empirical evidence on the cross-correlation between dimensional/ geometrical deviations and deviations from perfect function. This evidence is gathered always by measurement: tolerances so defined have some degree of uncertainty due to measurement errors. This leads to the application of worst-case tolerances, which are tighter than necessary to fulfil the allowable function variation.

Thus, the elimination of the zero-defect requirement, replacing full limit displacements by smaller ones, could make use of those extra allowances that are not in the drawing of the part, but are inherent in the tolerance allocation process. It is highly probable that the units accepted under this inspection condition will function properly. In addition, the production economy will result improved, because of the reduction of internal failure costs.

The selection and application of measurement systems to operate under these conditions still follows the recommendations formulated above. However, the requirement on measurement uncertainty can be slightly eased, because the diminution of limit displacements makes the value of D to decrease. Limit displacements should be assigned following the scheme in case study #1, *i.e.*, achieving a compromise between the excess of tolerance produced and the value of D (see §6.2.3).

The discussion about zero-defect is of no sense when talking about the application of measurement systems in classification tasks. Though limit displacements can be used in dimensional classification, its objective is not the same as in the inspection of batches manufactured to fulfil one- or two-sided tolerances. Indeed, limit displacements can be used to select the "kind of contamination" that each class suffers because of measurement errors and rounding effects. For example, the functional performance of a needle bearing can be strongly affected by a needle whose diameter is bigger than acceptable. Then, a displacement to the left could be applied in the upper class limits, reducing the probability to accept needles belonging to the class to the right. However, if the same concept is applied in all classes, the lower limit of the considered class will also be displaced to the left, resulting in the contamination of the class with undersized needles, that could not be so critical. The implementation of this type of solution depends on functional considerations that are beyond the scope of this thesis.

It has been already mentioned that measurement systems used in dimensional classification should be particularly accurate to minimise the fraction misclassified. However, this requirement is not easy to fulfil, because this kind of inspection is associated with small tolerance values, which can not be satisfied by the available manufacturing processes. For example, in the bearing industry it is not uncommon to separate the rolling elements in classes of $1\mu m$ width. Being the tolerances so tight, it is difficult to obtain measuring instruments with an uncertainty small enough to classify the parts efficiently. This situation results in scarcely acceptable values of the fraction misclassified. Nevertheless, the values of the D-measure are not as high as expected (see case study in §6.3). This can be attributed to the high value of the non-dimensional quality loss that is inherent in classification operations. If the fraction misclassified has to be reduced, the only way is to reduce measurement uncertainty by periodical correction of systematic errors and strict control of all the quantities affecting repeatability.

This section has been devoted to draw some basic rules for the selection and application of measurement systems in 100% inspection. The next one will discuss briefly the operational advantages of evaluating the inspection performance by simulation.

6.5 The evaluation of inspection performance by simulation

The discussion in the previous section showed that the simulation program proposed in this thesis is more than a tool to evaluate measurement system capability: it is a method to evaluate and improve the quality of batches subjected to 100% inspection. It provides information about the quality loss due to the complete production system (manufacturing and inspection), the fraction of this loss that can be attributed to the inspection system and the degree of contamination of the outgoing batch with units out of specification.

It has been shown that the algorithm emulates the behaviour of a complex system, with a high number of parameters that interrelate non-linearly. This complexity is hidden to the user, who only has to feed into the program the required data and interpret the results. Data are available in the shop floor and familiar to a quality technician or engineer with basic metrology knowledge. Results are meaningful in the industrial environment: the performance measures are in monetary units. They can be used to decide short-term actions on the process, but also to decide investments (e.g. improving manufacturing process capability to abandon 100% inspection).

An important advantage of the simulation approach is the capability of exploring a wide range of parameter combinations in a very short time interval. For example, all the data associated with case study #1 have been generated in approximately one hour (the platform used was a Pentium 200MHz-32Mb RAM microcomputer, which required approximately 20 seconds per complete simulation run). This characteristic encourages the appraiser to test different inspection conditions, so increasing the probability of finding an optimum solution to the problem. In addition, the appraiser learns about how inspection systems behave, by his or her interaction with the simulation program.

Finally, the use of simulation makes possible to predict the inspection performance based on *a priori* data. This permits improving the selection and design of inspection systems, embodying quality from the beginnings, saving money and development time.

The program described in this thesis is far from becoming a commercial product. Protection routines, error-handling routines, help and other professional attributes are lacking. Also the graphic capability need to be improved, perhaps by linking *Wininspect* with other commercial packages for data processing and visualisation. Nevertheless, the prototype program has succeeded in showing that simulation can be applied to solve concrete quality problems in industry.

7 BRINGING TOGETHER AND CONCLUSIONS

The following contributions have been made in this thesis:

- a new measure of the inspection performance has been proposed, based on the effect of measurement inaccuracy on the quality of manufactured units (chapter 3);
- a computer algorithm to simulate inspection operations has been developed, to evaluate the measure in practical situations (chapter 4);
- the behaviour of the proposed measure has been studied in different measurement and manufacturing conditions, comparing it also with the behaviour of other measures of inspection performance (chapter 5);
- a software prototype has been proposed to evaluate the performance of 100% inspection systems in industry (§6.1).
- Recommendations have been drawn to guide the selection and application of measurement systems in 100% inspection tasks (§6.4).

The measure proposed in this thesis, called *D*, evaluates the fraction of the non-dimensional quality loss per unit that can be attributed to inspection errors. It has been defmed using the quadratic quality loss function concept, modified to fit 100% inspection requirements. In consequence, *D* is a specific capability measure expressed in terms of monetary losses of product quality. Equations have been proposed to compute the value of *D* in several inspection cases: two-sided tolerances, one-sided tolerances and dimensional classification. Nominal-the-best, asymmetric and smaller-the-better quality loss functions have been used, depending on the inspection case.

Considering the difficulties of evaluating the inspection performance in industrial situations, the use of computer simulation has been proposed. The algorithm is based on the assumption that inspection errors are caused by the lack of measurement accuracy. Coarse operator errors and behavioural patterns have not been considered (in companies having developed quality systems, these errors are controlled by inspection procedures and operator training). The stochastic model of measurement process has been constructed according the concept of measurement uncertainty. Unknown and residual systematic errors have been treated as curves determined by random parameters, instead of independent random variables. In this way, the correlation of these errors with manufactured dimensions is respected, maintaining also its nature of uncertainty contributions. The proposed model considers that measurement and manufacturing processes can be non-gaussian. For those cases, it is suggested to apply the beta probability density function, as well as other probability models like triangular and rectangular.

The behaviour of D has been studied by computer simulation for gaussian manufactured populations subjected to two-sided tolerances. It has been shown that, even for a fixed inspection condition, the values of D scatter. The main reason of this scatter is the lack of knowledge of the value of systematic errors in the neighbourhood of each specification limit. It has been also shown that the value of D grows drastically with the application of limit displacements to preserve the quality of the outgoing units. In some cases, defmed by perfectly feasible combinations of parameters, up to 60% of the quality loss per unit could be attributed to the influence of metrology-related decisions. This can be traced to the massive rejection of conforming units due to the adoption of this practice.

In the field of capability evaluation, it has been demonstrated that measurement systems with the same uncertainty can result in different quality losses, depending on the inspection condition. In many cases the capability of manufacturing process and the value of limit displacements are more influent than measurement system properties. The same behaviour has been found when studying the cross-correlation between D and the direct measures of inspection performance described in §2.1: there is always some scatter. Then, it can be concluded that the D-measure defines a scale of inspection performance that is not consistent with other existing capability criteria. Nevertheless, the concept embodied in D is more meaningful in production environment. It does not evaluate the accuracy of measurement or the accuracy of actions on the product: it evaluates directly the increment in the quality loss due to non-ideal inspection.

The prototype software proposed in chapter 6 has been designed as a tool for quality improvement when 100% inspection is included in the process flowchart. It emphasises the multidisciplinary nature of quality planning, because requires the use of data coming from design, manufacturing and metrology. The results permit the optimisation of inspection conditions regarding the total quality loss, the fraction of quality loss due to non-ideal inspection and the degree of contamination of the accepted batch. This way, zero-defect or minimum cost criterions can be implemented, as well as other intermediate situations.

The application of the simulation software to an industrial case study shows that zero-defect can be only achieved by limit displacements of at least one measurement uncertainty. This manufacturing policy results in high internal failure costs when 100% inspection is used to preserve the quality of the product. The use of smaller limit displacements can improve production economy, if the customer accepts a slight extension of tolerance interval. Given the present situation in tolerance allocation, this tolerance extension would not be harmful for function in most cases. Thus, when 100% inspection is required, the size of limit displacements is a key issue in customer-supplier negotiation. The program proposed in this thesis could help in the negotiation. Nevertheless, it needs to be first validated and accepted by the involved parties or even by the standards community.

It becomes clear that, whenever possible, special and critic characteristics should be manufactured by process having enough capability. This would permit a more economic production and the application of statistical process control within the policy of continuous improvement. The proposal in this thesis could provide the information needed to decide weather or not an investment to improve manufacturing process capability can be justified by the diminution of internal and external failure costs.

A marginal benefit of the program proposed in chapter 6 is its eventual application when teaching industrial metrology at an undergraduate and graduate level. In the opinion of the author, it can be used to improve the perception of the multidisciplinary nature of quality assurance and its implications in manufacturing economy. In addition, the possibility to experiment with several combinations of parameters in short time intervals would allow perceiving the relevancy of quality control process optimisation.

It should be noted that this proposal is part of a higher-level objective: the development of a complete simulation environment to support measurement system selection and application during the advanced product quality planning process. To achieve this objective, it is natural to continue in the direction of statistically controlled processes. The main challenge in this case seems to be the simulation of out of control situations. However, this is part of a future research.

8 REFERENCES

/1/ C. Suntag, "Inspection and Inspection Management". ASQC Quality Press, Milwakee, 1993.

/2/ T. Pfeifer, "Qualitätsmanagement – Strategien, Methoden, Techniken". Carl Hanser Verlag, München, 1996. *(in German)*

/3/ J.M. Juran, F.M. Gryna, "Quality control handbook". 4th Ed., McGraw-Hill Book Co., New York, 1988.

/4/ L. Walsh, "Back to one hundred percent inspection?". Quality Management and Engineering, March 1974.

/5/ J. Peters, K.U. Leuven, "Metrology in design and manufacturing - facts and trends". Annals of the CIRP, Vol.26/2, pp.415-421,1977.

/6/ F.T. Farago, M.A. Curtiss, "Handbook of dimensional measurement". 3rd Ed., Industrial Press Inc., New York, 1994.

/7/ E.S. Roth, "Gages reject too high a percentage of good parts". Functional Gaging and Inspection Techniques Program, Soc. of Manufacturing Eng., Dearborn, 1983.

/8/ T. Pfeifer, "Fertigungmeßtechnik I". Lehrstuhl für Fertigungmeßtechnik und Qualitätsmanagement, WZL - RWTH Aachen, ca. 1995. *(in German)*

/9/ P. Bennich, "Dimensional quality control and quality management in manufacturing companies". Proc. of the III Int. Seminar on Metrology for Quality Control, Florianopolis, 1997.

/10/ BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, "International vocabulary of basic and general terms in metrology". DIN, Deutsches Institüt für Normung, Berlin, 1994.

/11/ ISO 9001, "Quality systems - model for quality assurance in design / development, production, installation and servicing". International Organization for Standardization, Geneva, 1994.

/12/ ISO 10012-1, "Quality assurance requirements for measuring equipment - Part 1: Metrological confirmation system for measuring equipment". International Organization for Standardization, Geneva, 1992.

/13/ G.D. Donatelli, "Selección de Sistemas de Medición para Tareas de Control de Calidad". Course offered during the I International Seminar on Metrology for Quality Control, Florianópolis, 1995. *(in Spanish)*

/14/ A. Tricker, E. Coates, E. Okell, "The effect on the R-chart of precision of measurement". J. of Quality Technology, Vol.30, No.3, pp.232-239, 1998.

/15/ H.-J. Mittag, D. Stemann, "Gauge imprecision effect on the performance of the $\overline{X}-S$ control chart". J. of Applied Statistics, Vol.25, No.3, pp.307-317, 1998.

/16/ H.-J. Mittag, "Measurement error effects on the performance of process capability indices". *In* H.-J. Lenz and P.-Th. Wilrich (Eds.), "Statistical quality control, Vol.5, pp.195-206, Physica, Heidelberg, 1997.

/17/ M. Hernla, "Meßunsicherheit und Fähigkeit". Qualität und Zuverlässigkeit, Vol.41, pp.1156-1162, 1996. (*in German*)

/18/ A. Weckenmann, "Causes and consequences of measurement uncertainty in production metrology". Proc. of the 6th Int. Symp. on Metrology for Quality Control - IMEKO, Vienna, pp.693-698, 1998.

/19/ G.D. Donatelli, A.M. Barp, C.A. Schneider, "The influence of measurement errors on the values of process capability indices". Proc. of the 6th Int. Symp. on Metrology for Quality Control - IMEKO, pp.133-141, Vienna, 1998.

/20/ L. Bauer, "Inspektionsfehler in der attributiven Qualitätaskontrolle". Arbeiten zur Angewandten Statistik, Physica-Verlag, Heidelberg, 1987. (*in German*)

/21/ Donatelli GD, Schneider CA, "Aptitud de los sistemas de medición para la verificación de tolerancias dimensionales". Proc. of the XIII Brazilian Congress on Mechanical Engineering, Belo Horizonte, 1995. *(in Spanish)*

/22/ W. Kessel, "Uncertainty of measurement: statement of physical ignorance". Paper presented at the Advanced School of Metrology: Uncertainty of Measurement, INMETRO, Angra dos Reis, 1997.

/23/ BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, "Guide to the expression of uncertainty in measurement". International Organization for Standardization, Geneva, 1993.

/24/ Kaarls R, "The treatment and statement of measurement uncertainty". Proc. of the 2nd. Symp. of the IMEKO Tech. Comitee in Metrology - Theoretical and Practical Limits of Measurement Accuracy", pp.21-30, Budapest, Romania, 1983.

/25/ NIS 3003, "The expression of uncertainty and confidence in measurement for calibration". NAMAS Executive - National Physical Laboratory, London, 1995.

/26/ O. Mathiesen, "Uncertainty of measurement - understanding the GUM". Paper presented at the Advanced School of Metrology: Uncertainty of Measurement, INMETRO, Angra dos Reis, 1997.

/27/ Berndt G, Hultzsch E, Weinhold H, "La tolerancia de función y la incertidumbre de medidas". J. of Jena, pp.263-270, 1968. (*in Spanish*)

/28/ ISO/FDIS 14253-1(E), "Geometrical product specifications (GPS) – Inspection by measurement of workpieces and measuring equipment. Part 1: Decision rules for proving conformance or non-conformance with specification". International Organization for Standardization, Geneva, 1997.

/29/ ISO/FDIS 14253-2(E), "Geometrical product specifications (GPS) – Inspection by measurement of workpieces and measuring equipment. Part 2: Guide to the estimation of uncertainty in GPS measurement, in calibration of measuring equipment and in product verification". International Organization for Standardization, Geneva, 1997.

/30/ Bosch Gmbh, "Quality assurance in the Bosch group No. 10 – Technical statistics: Capability of measuring facilities". Bosch Gmbh, Stuttgart, ca.1995.

/31/ AIAG, ASQC, "Measurement system analysis - Reference manual". Automotive Industries Action Group, Troy, 1995.

/32/ M. Hernla, "Prüfmittelfähigkeit is unzureichend". Qualität und Zuverlässigkeit, Vol.43, pp.194-196, 1998. *(in German)*

/33/ E. Dietrich, "Measurement system capability study – State of the art". Proc. of the 6th Int. Symp. on Metrology for Quality Control - IMEKO, Vienna, pp.125-132, 1998.

/34/ A. Weckenmann, *personal communication*. Vienna, 1998.

/35/ J.M. Juran, "Inspector's errors in quality control". Mechanical Engineering, Vol.59, No.10, pp.643-644, 1935.

/36/ H.H. Freeman, M. Friedman, F. Mosteller, W.A. Wallis, "Sampling Inspection". McGraw-Hill Book Co., New York, 1948.

/37/ R.L. McCornack, "Inspector accuracy: a study of the literature". Report SC-TM-610053-14, Sandia Corp., Albuquerque, 1961.

/38/ P.M. Wallack, S.K. Adams, "A comparison of inspector performance measures". AIIE Transactions, No.2, pp.97-105, 1970.

/39/ W.J. Youden, "Index for rating diagnostic tests". Cancer, No.3, pp.32-35, 1959.

/40/ L.S. Nelson, "Rating and discrimination of pass-fail tests". J. of Quality Technology, Vol.6, pp.58-59, 1974.

/41/ P.J. Tiemstra, "Measuring the performance of inspection and sorting systems". J. of Quality Technology, No.13, pp.149-156, 1981.

/42/ D.M. Green, J.A. Sweets, "Signal Detection Theory and Psychophysics". John Wiley & Sons, New York, 1966.

/43/ L.A. Goodman, W.H. Kruskal, "Measure of association for cross-classifications". J. of the American Statistical Association, 49, pp.732-764, 1954.

/44/ J. Cohen, "A coefficient of agreement for nominal scales". Educational and Psychological Measurement, No.20, pp.37-46, 1960.

/45/ Y.M.M. Bishop, S.E. Feinberg, P.W. Holland, "Discrete Multivariate Analysis". The MIT Press, Cambridge, 1975.

/46/ T. Raz, J.M. Liittschwager, "Ten measures of inspection performance". J. of Quality Technology, Vol.21, No.2, pp.81-86, 1989.

/47/ G.D. Donatelli, Schneider CA, "Classification performance of dimensional quality control systems: evaluating the influence of measurement errors". Proc. of the IMEKO TC-14 Int. Symp. on Metrology for Quality Control, pp.94-103, Zaragoza, 1995.

/48/ G. Taguchi, E.A. Elasayed, T.C. Hsiang, "Quality Engineering in Production Systems". McGraw-Hill, New York, 1989.

/49/ M.S. Phadke, "Quality Engineering Using Robust Design". PTR-Prentice Hall, Englewood Cliffs, 1989.

/50/ J.D.T. Tannock, C.F. Earl, "Assessment of quality performance in manufacturing". Total Quality Management, Vol.1, No.2, pp.197-206, 1990.

/51/ Ø. Bjørke, "Computer-Aided Tolerancing". 2nd Ed., ASME Press, New York, 1989.

/52/ E.I. Tsvetkov, "Simulation and modelling in measurement". Measurement, Vol.1, No.3, pp.129-132, 1983.

/53/ Black, M., "Modelos y metáforas". Ed. Tecnos, Madrid (edition and year not available).

/54/ Physicalisch-Technische Bundesanstalt, Labor 5.32, Megakal:meßaufgabenspezifische Kalibrierung von Koordinatenmeßgeräten durch Simulation (version 1.0).

/55/ J.B. Bryan, E.R. McClure, W. Brewer, J.W. Pearson, "Thermal effects in dimensional metrology". Proc. of the Metals Eng. and Production Eng. Conference - ASME, Berkeley, 1965.

/56/ J.B. Bryan, D.L. Carter, R.W. Clouser, J.H. Hamilton, "An order of magnitude improvement in thermal stability with use of liquid shower on a general purpose measuring machine". Proc. of the SME Precision Machining Seminar, 1982.

/57/ J.F. Minner, "Factor heat into your measurement". Quality Magazine, April 1996.

/58/ J.B. Bryan , "The deterministic approach in metrology and manufacturing". Proc. of the Int. Forum on Dimensional Tolerancing and Metrology, ASME Board on Research and Technology Development, Vol.27, pp.85-95, 1993.

/59/ G.D. Donatelli, M. Golfan, C.A. Schneider, "Measurement-based dimensional control systems: evaluating the influence of measurement errors and resolution". Proc. of the Int. Manufacturing Engineering Conference, Storrs, 1996.

/60/ A. Papoulis, "Probability, Random Variables and Stochastic Processes". McGraw-Hill, New York, 1965.

/61/ AIAG, ASQC, "Statistical Process Control - Reference manual". Automotive Industries Action Group, Troy, 1995.

/62/ E.M. Mansoor, "The application of probability to tolerances used in engineering design". Proc. of the Institution of Mechanical Engineers, No.178, pp.29-51, 1963/4.

/63/ J.R. He , "Estimating the distribution of manufactured dimensions with the beta probability density function". Int. J. of Machine Tool and Manufacture, Vol.31, No.3, pp.383-396, 1991.

/64/ P.Bennich, *personal communication*. Florianópolis, 1997.

/65/ S-S. Lin, H-P. Wang, C. Zhang, "Statistical tolerance analysis based on beta distributions". J. of Manufacturing Systems, Vol.16, No.2, pp.150-158, 1997.

/66/ S. Kotz, C.R. Lovelace, "Process Capability Indices in Theory and Practice". Arnold, London, 1998.

/67/ M. Golfan, C.A. Schneider, G.D. Donatelli, "Geração de números aleatórios com distribuição beta para simulação de processos metrológicos". Proc. of the III Congress of Mechanical Engineering Students, Rio de Janeiro, 1996. *(in Portuguese)*

/68/ E.A. Giesel, G.D. Donatelli, C.A. Schneider, "Uma rotina para simulação de erros de medição no cálculo de incertezas". Proc. of the IV Congress of Mechanical Engineering Students, Florianópolis, 1997. *(in Portuguese)*

/69/ K.V. Bury, "Statistical Models in Applied Science". J. Wiley & Sons, New York, 1975.

/70/ J. Mandel, T.W. Lashof, "The nature of repeatability and reproducibility". J. of Quality Technology, Vol.19, No.1, pp.29-36, 1987.

/71/ R.B.F. Schumacher, "Systematic Measurement Errors". J. of Quality Technology, Vol.13, No.1, pp.10-24,1981.

/72/ J. Mandel, "Some thoughts on variable selection in multiple regression". J. of Quality Technology, Vol.21, No.1, pp.2-6, 1989.

/73/ Y. Tomita, S. Honda, "Experimental evaluation of instrument scale readings obtained by eye interpolation of fractional interval values". Measurement 13, pp.147-151, 1994.

/74/ J. Banks, D. Goldsman, J.S.Carson, "Computer simulation": *in* H.M. Woodsworth (Ed.), "Statistical Methods for Engineers and Scientists". McGraw-Hill, New York, 1990.

/75/ P. Cooke, "Statistical inference for bounds of random variables". Biometrika, No.66, pp.367-374, 1969.

/76/ E. Kreyszig, "Introductory Mathematical Statistic – Principles and Methods". J. Wiley & Sons, New York, 1970.

/77/ Microsoft Developer Studio - Fortran Power Station 4.0. Microsoft Co., 1994-1995.

/78/ ISO 8015, "Technical drawings - Fundamental tolerancing principle". International Organization for Standardization, Geneva, 1985.

/79/ ISO 1101, "Technical drawings - Geometrical tolerancing - Tolerancing of form, orientation, position and run-out - Generalities, definitions, symbols, indications on drawings". International Organization for Standardization, Geneva, 1983.

/80/ G.D. Donatelli, C.A. Schneider, "Optimisation of Dimensional Quality Control: a Simulation Approach". Proc. of the III Int. Seminar on Metrology for Quality Control, Florianopolis, 1997.

/81/ D.J. Wheeler, R.W. Lyday, "Evaluating the Measurement Process". Statistical Process Control Press, Knoxville, 1984.